



# ***STIC Search Report***

## ***Biotech-Chem Library***

**STIC Database Tracking Number: 165311**

**TO: Medina A Ibrahim**  
**Location: REM/2B07/2C18**  
**Art Unit: 1638**

*Sept 14*, 2005

**Case Serial Number: 10/780002**

**From: P. Sheppard**  
**Location: Remsen Building**  
**Phone: (571) 272-2529**

**sheppard@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

From: Ibrahim, Medina A.  
Sent: Friday, September 09, 2005 12:30 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/780,002

Please search the following:

1. SEQ ID NO: 1 and 36
2. DNA encoding SEQ ID NO: 2.
3. an oligo of SEQ ID NO: 1 or 36.

Please search both commercial and patent (including pending) databases. Thanks

Medina A. Ibrahim  
Patent Examiner, GAU-1638  
transgenic plants and plant breeding  
Remsen-2B07  
mailbox-2C18  
(571) 272-0797

RECEIVED  
SEP - 9 2005  
STIC-BIOTECH/CHEM. DIVISION  
(STIC)

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**



```

Db      35  GGTTCGATCAGTGAAGAGATTACATTGTGTGCAC 71
|||||
RESULT 3
AAQ61367/c
ID  AAQ61367 standard; DNA; 258 BP.
XX
XX  AAQ61367;
AC
XX
XX  25-MAR-2003 (revised)
DT
XX  16-MAR-1994 (first entry)
DT
XX
XX  Human brain Expressed Sequence Tag EST01360.
DE
XX
XX  Gene transcription product; genetic markers; tagging; in vivo;
KW  transcription; mapping; locations; chromosomes; chromosomal; ss.
KW
XX
XX  Homo sapiens.
OS
XX
XX  W09316178-A2.
PN
XX
XX  19-AUG-1993.
PD
XX
XX  12-FEB-1993; 93WO-US001294.
PF
XX
XX  12-FEB-1992; 92US-00837195.
PR
XX
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA
XX
XX  Venter CU, Adams MD, Moreno RF;
PI
XX
XX  WPI; 1993-272882/34.
DR
XX
XX  Enriched oligonucleotides and corresp. sequences - used as markers for
PT  human genes transcribed in-vivo, facilitate tagging of most human genes.
PT
XX
XX  Example 4; Page 480; 500pp; English.
XX
XX  The Expressed Sequence Tag was isolated from a human brain cDNA library
CC  as part of a large set of ESTs which can be used as markers for human
CC  genes transcribed in vivo. They can be used to facilitate tagging of most
CC  human genes, for mapping locations of expressed genes on chromosomes, for
CC  individual or forensic identification, for mapping locations of disease-
CC  associated genes, for identification of tissue type, and for prepn. of
CC  antisense sequences, probes and constructs. EST01360 has a "poor" coding
CC  probability as evaluated using the coding-region prediction program CRM.
CC  See also AAQ59041-Q61440. (Updated on 25-MAR-2003 to correct PN field.)
CC
XX
XX  Sequence 258 BP; 54 A; 58 C; 75 G; 68 T; 0 U; 3 Other;
SQ
Query Match 2.5%; Score 27; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1049 CTTTCCTTTTGAAGAAAAA 1075
DB 136 CTTTCCTTTTGAAGAAAAA 110
RESULT 4
ABX54351/c
ID  ABX54351 standard; cDNA; 321 BP.
XX
XX  ABX54351;
AC
XX
XX  25-FEB-2003 (first entry)
DT
XX
XX  Bovine EST associated with lactation/muscle/fat deposition #4280.
DE
XX
XX  Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW  muscle deposition; fat deposition; genome mapping; gene identification;
KW

```

---

```

KW  Gene analysis; cattle breeding.
XX
XX  Bos Taurus.
XX
XX  US2002137160-A1.
PN
XX
XX  26-SEP-2002.
PD
XX
XX
XX  26-OCT-2001; 2001US-00983965.
PF
XX
XX  17-DEC-1998; 98US-0113678P.
PR
XX  15-DEC-1999; 99US-00465231.
PR
XX
XX  (BYAT/) BYATT J C.
PA  (MATH/) MATHIALAGAN N.
PA  (TAON/) TAO N.
PA  (WARR/) WARREN W C.
XX
XX  Byatt JC, Mathialagan N, Tao N, Warren WC;
PI  WPI; 2003-102386/09.
XX
XX  Purified nucleic acid molecules, useful for genome mapping, gene
PT  identification and analysis, cattle breeding or preparation of constructs
PT  for cattle gene expression and genetically improved cattle.
XX
XX  Claim 2; SEQ ID NO 4280; 38pp; English.
XX
XX  The invention relates to a purified nucleic acid molecule associated with
CC  lactation or muscle and fat deposition (designated LMFD), derived from
CC  cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC  nucleic acid molecule comprising any of 5912 nucleotide sequences,
CC  appearing as ABX50072-ABX55983, or complements of them. Also included are
CC  : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC  acid linked to a promoter and a 3' non-translated sequence that
CC  functions in the cell to cause termination of transcription and addition
CC  of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC  (2) determining a level or pattern of a molecule in a bovine cell or
CC  tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC  of the 5912 nucleic acid sequences or its complement or fragment) with a
CC  complementary nucleic acid molecule obtained from the bovine cell or
CC  tissue, where hybridisation between the marker nucleic acid and the
CC  complementary nucleic acid permits the detection of the molecule; and (b)
CC  detecting the level or pattern of the complementary nucleic acid, where
CC  the detection of the complementary nucleic acid is predictive of the
CC  level or pattern of the molecule. The LMFD nucleic acid is used for
CC  determining a level or pattern of a molecule in a bovine cell or tissue.
CC  It is useful for genome mapping, gene identification and analysis, cattle
CC  breeding, preparation of constructs for use in cattle gene expression, or
CC  for genetically improving cattle. The present sequence is one of the 5912
CC  bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present
CC  sequence was not shown in the specification but was obtained in
CC  electronic format from the USPTO web site:
CC  seqdata.uspto.gov/sequence.html?DocID=20020137160
XX
XX  Sequence 321 BP; 92 A; 72 C; 57 G; 100 T; 0 U; 0 Other;
SQ
Query Match 2.5%; Score 27; DB 8; Length 321;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1053 CCTTTTGAAGAAAAA 1079
DB 47 CCTTTTGAAGAAAAA 21
RESULT 5
ABL62609/c
ID  ABL62609 standard; DNA; 446 BP.
XX
XX  ABL62609;
AC
XX
XX  15-MAY-2002 (first entry)
DT

```

XX Colon adenocarcinoma related gene sequence SEQ ID NO:946.  
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX Homo sapiens.  
XX WO2001194629-A2.  
XX 13-DEC-2001.  
XX 30-MAY-2001; 2001WO-US010838.  
XX 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0233133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX Claim 1; SEQ ID NO 946; 44pp; English.  
XX The present invention describes a method (M1) for screening for an anti-

CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABU61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
CC tumour  
XX  
SQ Sequence 446 BP; 125 A; 97 C; 84 G; 140 T; 0 U; 0 Other;  
Query Match 2.5%; Score 27; DB 6; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1053 CCTTTTCAGAAAAA 1079  
Db 29 CCTTTTCAGAAAAA 3  
RESULT 6  
ABZ73132  
ID ABZ73132 standard; cDNA; 552 BP.  
XX AC ABZ73132;  
XX 10-APR-2003 (first entry)  
XX Rice leaf EST, SEQ ID NO:17.  
XX Rice; leaf; EST; expressed sequence tag; plant; biochip; DNA array;  
KW cloning; detection; heterosis; hybrid vigour; transgene detection;  
KW herbicide screening; pesticide screening; disease diagnosis;  
XX medical research; agriculture; ss.  
OS Oryza sativa.  
XX CN1364937-A.  
XX 21-AUG-2002.  
XX 31-OCT-2001; 2001CN-00137675.  
XX 31-OCT-2001; 2001CN-00137675.  
XX (UYZH-) UNIV ZHEJIANG.  
XX Li D, Dong H;  
XX WPI; 2003-000551/01.  
XX Rice leaf specific expression sequence label and its gene chip.  
XX Claim 1; Page 16 (Disclosure); 30pp; Chinese.  
XX The invention relates to 50 novel ESTs (expressed sequence tags; ABZ73116  
CC -ABZ73165) obtained from a rice leaf cDNA library. The invention also  
CC relates to a biochip comprising these ESTs. The biochip of the invention  
CC can be used in a variety of agricultural applications. It can be used in  
CC the cloning of genes which confer useful characteristics in crop species,  
CC to predict hybrid vigour (heterosis) at an early stage, in the detection  
CC of transgenic agricultural produce, in screening for novel herbicides and  
CC pesticides and in disease diagnosis. The biochip may also be used in  
CC medical research. The present sequence represents a rice leaf EST of the  
CC invention

```
XX SQ Sequence 552 BP; 135 A; 160 C; 105 G; 152 T; 0 U; 0 Other;
Query Match 2.5%; Score 27; DB 8; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCTTTTGAAGAAAAA 1077
Db 526 TTCTTTTGAAGAAAAA 552

RESULT 7
AAF16319
ID AAF16319 standard; cDNA; 1795 BP.
XX
AC AAF16319;
XX
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:754.
XX
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
XX Homo sapiens.
XX
XX WO200055174-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005988.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR P-PSDB; AAB57116.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
XX Claim 1; Page 1174-1175; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
XX Sequence 1795 BP; 473 A; 375 C; 447 G; 497 T; 0 U; 3 Other;

Query Match 2.5%; Score 27; DB 3; Length 1795;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1053 CCTTTTGAAGAAAAA 1079
Db 1757 CCTTTTGAAGAAAAA 1783

RESULT 8
ABL90691
ID ABL90691 standard; cDNA; 1795 BP.
XX
AC ABL90691;
XX
XX 24-MAY-2002 (first entry)
XX Human polynucleotide SEQ ID NO 1253.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
DR P-PSDB; ABB90282.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive, and
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
XX Claim 4; SEQ ID NO 1253; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1795 BP; 473 A; 376 C; 447 G; 497 T; 0 U; 2 Other;
SQ
Query Match 2.5%; Score 27; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079
Db 1757 CCTTTTGAAGAAAAA 1783
```

CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent  
 CC skin aging due to sunburn, to change a mammal's mental state or physical  
 CC state by influencing biorhythms, cardiac rhythms, depression, memory,  
 CC stress and for accelerating wound healing. (I), (II) and/or their agonist  
 CC or antagonist are useful as food additives or preservatives to increase  
 CC or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamin, mineral or other nutritional components. (I) is  
 CC useful for screening therapeutic compounds. (II) is useful in forensic  
 CC biology for detecting DNA sequences and as diagnostic probes for  
 CC detecting the presence of specific mRNA in a particular cell type  
 XX  
 SQ Sequence 342 BP; 116 A; 77 C; 59 G; 90 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 3; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAATAAAAAAAAAAAAAA 1079  
 DB 316 CTTTGTGAAAAAATAAAAAAAAAAAAAA 341

RESULT 10  
 AAC79873  
 ID AAC79873 standard; cDNA; 375 BP.  
 XX  
 AC AAC79873;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA for gene 25.  
 XX  
 KW Human; secreted protein; cytostatic; antiarthritic; antiasthmatic;  
 KW immunosuppressive; antiarteriosclerotic; antiinflammatory; nootropic;  
 KW neuroprotective; antidiabetic; tranquiliser; vulnary; antibacterial;  
 KW antipsoriatic; antiarrhythmic; antirheumatic; cardiant; anti-HIV;  
 KW autoimmune disorder; allergic condition; cardiovascular disorder; cancer;  
 KW neurological disease; tissue repair; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055176-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 09-MAR-2000; 2000WO-US006057.  
 XX  
 PR 12-MAR-1999; 99US-0124142P.  
 PR 11-JUN-1999; 99US-0138597P.  
 PR 03-DEC-1999; 99US-0168666P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 DR WPI: 2000-638176/61.  
 DR P-PSDB; AAB44860.  
 XX  
 PT Novel 49 human secreted proteins useful for diagnosis, prevention and  
 PT treatment of disorders including neurological, cell proliferative,  
 PT cardiovascular, and autoimmune/inflammatory disorders and microbial  
 PT infections.  
 XX  
 PS Claim 1a; Page 351; 405pp; English.

This invention describes a novel isolated polypeptide (I) comprising an  
 CC amino acid sequence at least 95 % identical to a polypeptide sequence  
 CC selected from 49 polypeptides encoded by polynucleotide sequences  
 CC included in American Type Culture Collection (ATCC) deposit number  
 CC 203917, defined in the specification. The products of the invention have  
 CC cytostatic, antiarthritic, antiasthmatic, immunosuppressive, nootropic,  
 CC antiarteriosclerotic, antiinflammatory, neuroprotective, antidiabetic,  
 CC tranquiliser, vulnary, antibacterial, antipsoriatic, antiarrhythmic,  
 CC antirheumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II)  
 CC encoding (I) is useful for preventing, treating or ameliorating a medical  
 CC condition and for diagnosing a pathological condition or susceptibility  
 CC to the condition. (I) is useful for identifying a binding partner which  
 CC affects the activity of the polypeptide and for identifying an activity  
 CC in a biological sample. (I), (II) or an antibody (IV) specific to (I) is  
 CC also useful for treating or preventing a disease, disorder or condition  
 CC associated with aberrant expression of (I). Diseases treated or diagnosed  
 CC include immune disorders such as autoimmune diseases, blood protein  
 CC disorders, anemia, allergic reactions and conditions such as asthma,  
 CC organ rejection or graft-versus-host disease, inflammation, hyper  
 CC proliferative disorders, cardiovascular disorders such as arterioarterial  
 CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ  
 CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy,  
 CC rheumatoid arthritis, psoriasis, diseases associated with increased  
 CC apoptosis that include acquired immunodeficiency syndrome (AIDS),  
 CC neurological diseases such as Parkinson's disease, viral, bacterial,  
 CC fungal or parasitic diseases. They are also used to repair, replace or  
 CC protect tissue damage by congenital defects, to treat trauma, in surgery,  
 CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury

CC antirheumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II) encoding (I) is useful for preventing, treating or ameliorating a medical condition and for diagnosing a pathological condition or susceptibility to the condition. (I) is useful for identifying a binding partner which affects the activity of the polypeptide and for identifying an activity in a biological sample. (I), (II) or an antibody (IV) specific to (I) is also useful for treating or preventing a disease, disorder or condition CC associated with aberrant expression of (I). Diseases treated or diagnosed CC include immune disorders such as autoimmune diseases, blood protein CC disorders, anemia, allergic reactions and conditions such as asthma, CC organ rejection or graft-versus-host disease, inflammation, hyper CC proliferative disorders, cardiovascular disorders such as arterioarterial CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy, CC rheumatoid arthritis, psoriasis, diseases associated with increased CC apoptosis that include acquired immunodeficiency syndrome (AIDS); CC neurological diseases such as Parkinson's disease, viral, bacterial, CC fungal or parasitic diseases. They are also used to repair, replace or CC protect tissue damage by congenital defects, to treat trauma, in surgery, CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent CC skin aging due to sunburn, to change a mammal's mental state or physical CC state by influencing biorhythms, cardiac rhythms, depression, memory, CC stress and for accelerating wound healing. (I), (II) and/or their agonist CC or antagonist are useful as food additives or preservatives to increase CC or decrease storage capabilities, fat content, lipid, protein, CC carbohydrate, vitamin, mineral or other nutritional components. (I) is CC useful for screening therapeutic compounds. (II) is useful in forensic CC biology for detecting DNA sequences and as diagnostic probes for CC detecting the presence of specific mRNA in a particular cell type

XX  
SQ Sequence 375 BP; 111 A; 89 C; 74 G; 96 T; 0 U; 5 Other;

Query Match 2.4%; Score 26; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAA 1079  
DB 299 CTTTGTGAAAAA 324

RESULT 11  
AAF26589  
ID AAF26589 standard; DNA; 375 BP.  
AC AAF26589;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE DNA encoding human secreted protein #43.  
XX  
KW Secreted protein; gene therapy; vaccine; cancer; leukemia;  
KW autoimmune disease; allergy; inflammation; graft rejection;  
KW hyperproliferation; cardiovascular; infection; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200076531-A1.  
XX  
PD 21-DEC-2000.  
XX  
PF 01-JUN-2000; 2000WO-US015137.  
XX  
PR 11-JUN-1999; 99US-0138625P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX  
DR WPI; 2001-071148/08.  
XX  
PT Nucleic acids encoding 47 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
PT and diabetic retinopathy.  
XX  
PS Claim 1; Page 469; 525pp; English.  
XX  
CC The present invention relates to 26 secreted human proteins. The proteins  
CC may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. For example, they  
CC may be used in gene therapy or in vaccines. Typical of diseases which are  
CC potentially treatable are cancers (including leukemia), autoimmune  
CC diseases, allergies, inflammation, graft rejection, hyperproliferation,  
CC cardiovascular diseases (particularly critical limb ischemia and coronary  
CC disease) and any involving abnormal angiogenesis, neurodegeneration  
CC and/or infectious diseases

XX  
SQ Sequence 375 BP; 111 A; 89 C; 74 G; 96 T; 0 U; 5 Other;

Query Match 2.4%; Score 26; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAA 1079  
DB 299 CTTTGTGAAAAA 324

RESULT 12  
AAS11362  
ID AAS11362 standard; DNA; 814 BP.  
XX  
AC AAS11362;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Soybean DNA encoding invertase inhibitor #2.  
XX  
KW Soybean; invertase inhibitor; ds; gene therapy; environmental stress;  
KW kernel development; antisense.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 99..638  
FT /\*tag= a  
FT /product= "Invertase inhibitor"  
FT /note= "This sequence is specifically claimed in claim 2"  
XX  
PN WO200158939-A2.  
XX  
PD 16-AUG-2001.  
XX  
PF 12-FEB-2001; 2001WO-US004492.  
XX  
PR 10-FEB-2000; 2000US-0181509P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
PA (DUPO) DU PONT DE NEMOURS & CO E I.  
XX  
PI Helentjaris T, Bate NJ, Allen SM;  
XX  
DR WPI; 2001-502706/55.  
DR P-PSDB; AAU05777.  
XX  
PT An isolated polypeptide when recombinantly expressed in a plant is useful  
PT for modulating invertase activity and increasing yield in the plant.  
XX  
PS Claim 2; Page 67-68; 83pp; English.  
XX  
CC The invention relates to novel plant invertase inhibitors or a yeast  
CC invertase which, when recombinantly expressed in a plant, can modulate  
CC invertase activity and increase yield in the plant. Chimeric invertase  
CC inhibitors are useful for modulating invertase activity and increasing  
CC yield in a plant (especially crop species) when used to transform the

CC plant and are also useful for modulating kernel development and  
 CC protecting plants against the harmful/detrimental effects of stress and  
 CC adverse environmental conditions. Yeast invertase is less sensitive to  
 CC invertase inhibitors therefore is an attractive option to supplement  
 CC invertase activity in a plant using gene therapy. The novel invertase  
 CC inhibitor nucleic acids may be used in their antisense form. The present  
 CC sequence encodes a soybean invertase inhibitor

XX Sequence 814 BP; 241 A; 183 C; 138 G; 252 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 5; Length 814;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAA 1079  
 DB 770 CTTTGTGAAAAA 795

RESULT 13

ADA13386  
 ID ADA13386 standard; cDNA; 1099 BP.

AC ADA13386;

DT 06-NOV-2003 (first entry)

XX Human intracellular signalling molecule INTSIG-23 cDNA, SEQ ID NO:68.

XX Human; intracellular signalling molecule; INTSIG;  
 KW cell proliferative disorder; cancer; atherosclerosis;  
 KW autoimmune disorder; inflammatory disorder; infection;  
 KW neurological disorder; developmental disorder; endocrine disorder;  
 KW cytoskeletal; antiarteriosclerotic; nootropic; neuroprotective;  
 KW cerebroprotective; immunosuppressive; antiinflammatory; gene therapy;  
 KW gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 98..625  
 FT CDS /\*tag= a  
 FT /product= "INTSIG-23"  
 FT

XX WO2003031568-A2.

XX 17-APR-2003.

XX 16-AUG-2002; 2002WO-US026322.

XX 17-AUG-2001; 2001US-0313245P.

XX 24-AUG-2001; 2001US-0314751P.

XX 31-AUG-2001; 2001US-0316752P.

XX 31-AUG-2001; 2001US-0316847P.

XX 14-SEP-2001; 2001US-0322188P.

XX 28-SEP-2001; 2001US-0326390P.

XX 12-OCT-2001; 2001US-0328952P.

XX 19-OCT-2001; 2001US-0345468P.

XX 12-APR-2002; 2002US-0372499P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lu DAM, Swannakar A, Tang YT, Griffin JA, Emerling BM;

XX Forsythe IJ, Yao MG, Ramkumar J, Richardson TW, Becha SD, Lee BA;

XX Warren BA, Lehr-Nason PM, Baughn MR, Li JX, Duggan BM, Gietzen KJ;

XX Lal PG, Borowsky ML, Ison CH, Thangavelu K, Xu Y, Lee S;

XX Elliott VS, Sprague WW, Azimzai Y, Hafalia AJA, Ding L, Nguyen DB;

PT diagnosing, treating and preventing diseases or conditions associated  
 PT with the aberrant INTSIG expression e.g. cancer, diabetes, epilepsy, or  
 PT infections.

XX Claim 5; Page 339; 367pp; English.

XX The invention relates to 45 human intracellular signalling molecules  
 CC (INTSIG), INTSIG-1 to INTSIG-45, and the cDNAs encoding them (ADA13319-  
 CC ADA13408). The invention also encompasses expression constructs, host  
 CC cells and transgenic organisms comprising an INTSIG nucleic acid sequence  
 CC ; the recombinant preparation of a INTSIG; an antibody against a INTSIG;  
 CC methods of detection of INTSIG proteins or nucleic acids; a micro-array  
 CC containing INTSIG nucleic acids; methods of screening compounds for their  
 CC ability to modulate INTSIG activity or expression; and pharmaceutical  
 CC compositions comprising an INTSIG protein, an INTSIG antibody, an INTSIG  
 CC agonist or INTSIG antagonist. The INTSIG proteins, nucleic acids or  
 CC compositions comprising them are useful in diagnosing, treating or  
 CC preventing a variety of disorders, including cell proliferative disorders  
 CC (e.g., cancer or atherosclerosis); autoimmune/inflammatory disorders  
 CC (e.g., AIDS, allergies, anaemia, asthma, bronchitis, gout, multiple  
 CC sclerosis, osteoarthritis, rheumatoid arthritis, or systemic lupus  
 CC erythematosus); neurological disorders (e.g., epilepsy, stroke,  
 CC Alzheimer's disease, dementia, or Parkinson's disease); viral, bacterial,  
 CC fungal, parasitic, protozoan or helminthic infections; neurological  
 CC disorders (e.g., epilepsy, stroke, Alzheimer's disease, dementia, or  
 CC Parkinson's disease); developmental disorders (e.g., achondroplastic  
 CC dwarfism, Cushing's disease, hypothyroidism or hydrocephalus); endocrine  
 CC disorders (e.g., disorders of the hypothalamus or pituitary gland or  
 CC diabetes); reproductive disorders; or vesicle-trafficking disorders  
 CC (e.g., cystic fibrosis or hypercholesterolaemia). The present sequence  
 CC represents an INTSIG cDNA of the invention.

XX Sequence 1099 BP; 397 A; 166 C; 227 G; 309 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 10; Length 1099;

Best Local Similarity 100.0%; Pred. No. 5.6;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAA 1079

DB 1069 CTTTGTGAAAAA 1094

RESULT 14

AAQ73389

ID AAQ73389 standard; cDNA to mRNA; 1664 BP.

XX AAQ73389;

XX 25-MAR-2003 (revised)

DT 09-JUN-1995 (first entry)

XX Rabies virus glycoprotein coding sequence.

XX Rabies virus; glycoprotein; cancer; vaccine; colonic cancer; melanoma;

XX breast cancer; kidney; tumour cell; fibroblast; transfect;

XX transformed cells; cytokine; co-stimulating molecule; culture;

XX heat stable antigen; ss.

XX Rabies virus.

XX Key Location/Qualifiers

FT 5'UTR 1..6

FT /\*tag= a

FT CDS 7..1581

FT /\*tag= b

FT sig\_peptide 7..63

FT /\*tag= c

FT mat\_peptide 64..1578

FT /\*tag= d

XX WO9421808-A1.

XX New human intracellular signalling molecules (INTSIG)), useful for



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 06:38:26 ; Search time 4681.81 Seconds  
(without alignments)  
11167.303 Million call updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaagaaaa.....gaaaaaagaaaaaa 1079

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1079	100.0	1079	8	AY485932	Nicotiana
2	37	3.4	121	6	CQ0809241	Sequence
C 3	30	2.8	2810	10	BC080732	Sequence
C 4	30	2.8	12336	3	AC116550	Dicystose
5	30	2.8	199593	10	AL671871	Mouse DNA
6	30	2.8	250078	3	AE014829	Plasmid
7	30	2.8	252632	3	AE014818	Plasmid
C 8	29	2.7	250029	3	AE014820	Plasmid
C 9	28	2.6	118189	2	AL513496	Homo sapi
C 10	28	2.6	145638	9	HS357K22	Human DNA
11	28	2.6	224697	2	AC135583	Rattus no
12	28	2.6	230038	2	AC109654	Rattus no
13	28	2.6	257686	2	AC121225	Rattus no
14	27	2.5	207	6	CQ672612	Sequence
C 15	27	2.5	446	6	AX330437	Sequence
16	27	2.5	1778	9	BC065299	Homo sapi
17	27	2.5	1830	9	BC063001	Homo sapi
C 18	27	2.5	138461	2	AC113576	Tetraodon
19	27	2.5	156881	5	BX088545	Zebrafish

20	2.5	162712	2	AC015677	Homo sapi
C 21	2.5	179391	2	AP001283	Homo sapi
C 22	2.5	206736	2	AP001200	Homo sapi
23	2.4	332	6	CQ709902	Sequence
24	2.4	814	6	AR491517	Sequence
25	2.4	814	6	AX214366	Sequence
26	2.4	1197	9	AB046068	Macaca fa
27	2.4	1234	9	AF038169	Homo sapi
28	2.4	1495	9	BC063688	Homo sapi
29	2.4	1622	5	BC061308	Xenopus t
30	2.4	1664	6	A40148	Sequence 1
31	2.4	1839	10	AY138960	Mus muscu
32	2.4	1860	10	BC027433	Mus muscu
33	2.4	1870	6	BD131320	Human rec
34	2.4	2013	8	AY735703	Arabidops
35	2.4	2080	8	PTV13772	Populus tri
36	2.4	2359	9	BC050626	Homo sapi
37	2.4	5966	9	HS0808857	Homo sapi
C 38	2.4	24739	3	AC114265	Dicystose
C 39	2.4	24739	3	AY621074	Homo sapi
40	2.4	41721	9	AC145291	Homo sapi
41	2.4	41823	9	AC092428	Homo sapi
C 42	2.4	67322	2	AC102007	Mus muscu
43	2.4	91692	9	AC084253	Homo sapi
44	2.4	98345	6	CQ870165	Sequence
C 45	2.4	113153	5	BX005344	Zebrafish

ALIGNMENTS

RESULT 1	AY485932	1079 bp	mRNA	linear	PLN 07-JAN-2004
LOCUS	Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.				
DEFINITION	AY485932				
ACCESSION	AY485932				
VERSION	AY485932.1				
KEYWORDS	Nicotiana tabacum (common tobacco)				
SOURCE	Nicotiana tabacum				
ORGANISM	Nicotiana tabacum				
REFERENCE	Kumar, D. and Klessig, D.F.				
AUTHORS	High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)				
JOURNAL	PUBMED				
REFERENCE	2 (bases 1 to 1079)				
AUTHORS	Kumar, D. and Klessig, D.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA				
FEATURES	Location/Qualifiers				
Source	1. 1079				
	/organism="Nicotiana tabacum"				
	/mol_type="mRNA"				
	/cultivar="Xanthi nc [NN]"				
	/db_xref="taxon:4097"				
	42. 824				
	/note="lipase; SAMP2"				
CDS	/codon_start=1				
	/product="salicylic acid-binding protein 2"				
	/protein_id="AA087711.1"				
	/db_xref="GI:40549303"				
	/translation="MKEGKHFLVHGACHGWSYKLPLEAAGHKVTLDAASGTDLRKTEELTYDYTLPLMEWESLSADEKVLVGHSLGNNLGLAMEKYPKIVAAVFLAAMPVSNSSFLVQYNERTEAENWDTQFLPYGSPPEPLSMFPGPKFLAHLVOLCSMPDALASSLVRPSSFLMEDLSKAKYFTDRFGSVKRVYIIVCTEDKRGIPREFQRWQIDNIGVTEAIEIKGDHMAMLCEPKQKLCASLLEIAHKYN"				
ORIGIN					



JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2810)  
Director MGC Project.  
Direct Submission  
Submitted (27-AUG-2004) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-s@mail.nih.gov](mailto:cgapbs-s@mail.nih.gov)  
Tissue Procurement: Dr. David Rowe  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nhgri.nih.gov](mailto:nisc.mgc@nhgri.nih.gov)  
Akhtar, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 177 Row: i Column: 22  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 31543884.

## FEATURES

Location/Qualifiers  
1. .2810  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="MGC:91232 IMAGE:6518356"  
/tissue\_type="Limb, mouse, day 10.5 and 11.5, pool of  
undifferentiated limb containing undifferentiated limb  
mesenchyme and early condensing mesenchyme."  
/clone\_lib="NIH MGC\_134"  
/lab\_host="DH10B"  
/notes="Vector: pCMV-SPORT6.1"  
1. .2810  
/genes="Tox"  
/db\_xref="LocusID:252838"  
/db\_xref="MGI:2181459"  
150. .1730  
/genes="Tox"  
/codon\_start=1  
/product="thymocyte selection-associated HMG box gene"  
/protein\_id="AAH80732.1"  
/db\_xref="GI:51593441"  
/db\_xref="LocusID:252838"  
/db\_xref="MGI:2181459"  
/translation="MDVRFYPPAPAAAPAAAPCLGSPCLDPYYCNKFDGNNYMSM  
TEPSQDYVASQSPGFSLESDFNIPPTPSLPDHSVLHNEVESGYRSLCHPMNH  
NGLLPFPOTMDLPEITVSNMLQDGLLSNSISVMQDIGNAEGAYSSHPOAAWRP  
RGQPTDRIQASMMQPGQLTTINQSLSAQLGMLMGCTVNAHNSPSPGSKSATSPS  
SSVHEDECEDASKINGEKRPASDMGKKPKPKPKKKKDPNEPQKPVSAIYALFFRDQTQ  
AAIKGNPNATFGEVSKI VASMMWDGIGBEQKVYKKKTEAAKKEYLKLQAAAYRSLVS  
KSTYDQVDTKTSQPLVNSKPSVHFQSPGSAHLSYLSHSHYQQQPMQLTAMHPSL  
PIAPKPNKMEVTSIANMAVSPPPPLQIGPPLHQLHLSMQQHQSLAMQPLGSQLP  
MOVOTALHSPTMQGFTLOPDYQTIINPTSTAAQVVTQAMEYVRSRCRNPFPQPDVDS  
TDYCSSGQGRDKALYI"

## gene

## CDS

Query Match  
Best Local Similarity 2.8%; Score 30; DB 10; Length 2810;  
100.0%; Pred. No. 0.00025;

## ORIGIN

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCTCTTTTGAAGAAAAA 12336 bp DNA linear INV 12-MAR-2003  
||||| 1079  
Db 119 TTTCTCTTTTGAAGAAAAA 100  
||||| 90

RESULT 4  
AC116550/c AC116550 12336 bp DNA linear INV 12-MAR-2003  
LOCUS Dictyostelium discoideum chromosome 2 map 375056-387392 strain AX4,  
DEFINITION complete sequence.  
ACCESSION AC116550  
VERSION AC116550.2 GI:28830294  
KEYWORDS HTG.  
SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE 1 (bases 1 to 12336)  
AUTHORS Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P.,  
Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K.,  
Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and  
Noegel, A.A.  
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum  
JOURNAL Nature 418 (6893), 79-85 (2002)  
MEDLINE 22092622  
PUBMED 12097910  
REMARK The Dictyostelium Genome Sequencing Consortium  
REFERENCE 2 (bases 1 to 12336)  
AUTHORS Baumgart, C.  
TITLE Direct Submission  
JOURNAL Submitted (29-MAR-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
REFERENCE 3 (bases 1 to 12336)  
AUTHORS Baumgart, C.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
REFERENCE 4 (bases 1 to 12336)  
AUTHORS Baumgart, C.  
TITLE Direct Submission  
JOURNAL Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
COMMENT On Mar 4, 2003 this sequence version replaced gi:19807753.  
CDS predictions from GeneID do not necessarily reflect true genes.  
Further information is available from IMB Jena, Department of  
Genome Analysis  
(<http://genome.imb-jena.de/dictyostelium/>)  
and the University Cologne, Institute for Biochemistry I  
(<http://www.uni-koeln.de/dictyostelium/project.shtml>)  
Funding  
Agency : Deutsche Forschungsgemeinschaft (DFG).  
Location/Qualifiers  
1. .12336  
/organism="Dictyostelium discoideum"  
/mol\_type="genomic DNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/chromosome="2"  
/map="375056-387392"  
692. .1432  
/note="GeneID exon scores (in order of location ranges) :  
51.49 - GSCI ID dd\_00837"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAO52649.1"  
/db\_xref="GI:28830295"  
/translation="MKLLSLILFTIAFCNGESLKGNNVDFSKLENKVGQOCFT  
SSSNFPTGYSTSFVQIGNANTPTNSNGDITOTSYMESGTVDIDFVGQLYVPFEMDTE  
NSVTGKLGWNGNNTQVNIENGNVEYCFLETLYSPFENLKYVSDAEGSVCECD  
LFMYPISILGNSTNQYLFVDKSDCLISASQANANGTPSGNSLTNFTYYQPSADSSNFQ  
LPSCILNGHISLSQPKLKLNLKPKIN"  
Join(2339..2348,2438..2598)

FEATURES  
source  
1. .12336  
/organism="Dictyostelium discoideum"  
/mol\_type="genomic DNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/chromosome="2"  
/map="375056-387392"  
692. .1432  
/note="GeneID exon scores (in order of location ranges) :  
51.49 - GSCI ID dd\_00837"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAO52649.1"  
/db\_xref="GI:28830295"  
/translation="MKLLSLILFTIAFCNGESLKGNNVDFSKLENKVGQOCFT  
SSSNFPTGYSTSFVQIGNANTPTNSNGDITOTSYMESGTVDIDFVGQLYVPFEMDTE  
NSVTGKLGWNGNNTQVNIENGNVEYCFLETLYSPFENLKYVSDAEGSVCECD  
LFMYPISILGNSTNQYLFVDKSDCLISASQANANGTPSGNSLTNFTYYQPSADSSNFQ  
LPSCILNGHISLSQPKLKLNLKPKIN"  
Join(2339..2348,2438..2598)

CDS

```
/note="GeneID exon scores (in order of location ranges):
0.99, 12.56 - GSCJ_ID dd_00838"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52650.1"
/db_xref="GI:28830296"
/translation="MLFKSIISLSNKSNNVSSSKLLSNDNSNLNSFDLTCNTVSNKSS
TTFFTRPNFVNY
3144. .3884
CDS
/note="GeneID exon scores (in order of location ranges):
60.79 - GSCJ_ID dd_00839"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52651.1"
/db_xref="GI:28830297"
/translation="MKLLLSLLFTIAFCGESLKLGVNNVDSKLENKVGQOQCF
TSSNFPPTGYSTFVQIAFNANTPTSGDITQTSYMSGTVDIDFVGGLYVDFEMDT
ENSVTKGLWGNNTQVNIELAGNEVCFLETITYSVPFENLKYVSDAIGSVEC
DLFMPSILGNSTNQTLFVDKSDCLISASAQNTGTGSGNSLTNFYIYEPGADSSNF
QLPSCILNSPILISQPKLSLLKRPKTL
join(4756. .4765,4837. .4997)
CDS
/note="GeneID exon scores (in order of location ranges):
0.99, 9.22 - GSCJ_ID dd_00840"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52652.1"
/db_xref="GI:28830298"
/translation="MLFKSIISLSNKSNNVSSSKLLSNDNSNLNSFDLTCNSVSNKSS
STFFTRPNFVNY
5618. .6358
CDS
/note="GeneID exon scores (in order of location ranges):
61.58 - GSCJ_ID dd_00841"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52653.1"
/db_xref="GI:28830299"
/translation="MKLLLSLLFTIAFCGESLKLGVNNVDSKLENKVGQOQCF
TSSNFPNGYSTFVQIAFNANTPTSGDITQTSYMSGTVDIDFVGGLYVDFEMDT
ENSVTKGLWGNNTQVNIELAGNEVCFLETITYSVPFENLKYVSDAIGSVEC
DLFMPSILGNSTNQTLFVDKSDCLISASAQNVGSPGNSLTNFYIYEPGADSSNF
QLPSCILNSPILISQPKLSLLNKIPKIN"
CDS
/note="GeneID exon scores (in order of location ranges):
61.85 - GSCJ_ID dd_00836"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52654.1"
/db_xref="GI:28830300"
/translation="MKLLLSLLFTIAFCGESLKLGVNNVDSKLENKVGQOQCF
TSSNFPPTGYSTFVQIAFNANTPTSGDITQTSYMSGTVDIDFVGGLYVDFEMDT
ENSVTKGLWGNNTQVNIELAGNEVCFLETITYSVPFENLKYVSDAIGSVEC
DLFMPSILGNSTNQTLFVDKSDCLISASAQNVGSPGNSLTNFYIYEPGADSSNF
QLPSCILNSPILISQPKLSLLNKIPKIN"
CDS
/note="GeneID exon scores (in order of location ranges):
9618. .9627, 9715. .9863, 9939. .9956"
/note="GeneID exon scores (in order of location ranges):
0.99, 9.22, -1.70 - GSCJ_ID dd_03027"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52655.1"
/db_xref="GI:28830301"
/translation="MLFKSIISLSNKSNNVSSSKLLSNDNSNLNSFDLTCNTVSNKSS
STFFTRPNFGPTNV"
CDS
/note="GeneID exon scores (in order of location ranges):
join(11040. .11049,11127. .11287)
/note="GeneID exon scores (in order of location ranges):
0.99, 11.60 - GSCJ_ID dd_03033"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO52656.1"
/db_xref="GI:28830302"
/translation="MLFKSIISLSNKSNNVSSSKLLSNDNSNLNSFDLTCNTVSNKSS
TTFFTRPNFVNY
complement(11510. .12193)
/note="GeneID exon scores (in order of location ranges):
```

```
61.72 - GSCJ_ID dd_03010"
/codon_start=1
/product="similar to Dictyostelium discoideum (Sline
mold). Phosphatidylinositol phosphate kinase 6 (Fragment)"
/protein_id="AAO52657.1"
/db_xref="GI:28830303"
/translation="MTTLFDLTLNNTNNNNNNYAGCKRQHSINKRVDLIIPSDVTL
TNDKLLIETELTGSKNDIDINIKSDILIIQGERKKKSIIRHQQQQHQHQQQQQQQQ
QQQQQQQQQQQQQQQQQQQLENSKENDEPSIEFEDVKSELNKTLTNTTENKD
EDKTTQISKKFISERSFGNSKRYLNLSEILYQLDLNSINTQFENGLLTIITKKKFD
SNTNIKININ"
QY 1050 TTTCTCTTTTTCAGAAAAA 1079
Db 10343 TTTCTCTTTTTCAGAAAAA 10314
ORIGIN
Query Match 2.8%; Score 30; DB 3; Length 12336;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1050 TTTCTCTTTTTCAGAAAAA 1079
Db 10343 TTTCTCTTTTTCAGAAAAA 10314
RESULT 5
AL671871
LOCUS Mouse DNA sequence from clone RP23-157M14 on chromosome 4, complete
sequence.
ACCESSION AL671871
VERSION AL671871.4 GI:21212208
KEYWORDS HTG
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 199593)
AUTHORS Tracey A.
JOURNAL Direct Submission
COMMENT Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:201522695.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACes.6.
Location/Qualifiers
1. .199593
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-157M14"
/clone_lib="RPCI-23"
ORIGIN
```

```
Query Match          2.8%; Score 30; DB 10; Length 199593;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1050 TTTCCTTTTGAAGAAAAAAGAAAAAAGAAAAA 1079
Db      166968 TTTCCTTTTGAAGAAAAAAGAAAAAAGAAAAA 166997

RESULT 6
AE014829
LOCUS   Plasmodium falciparum 3D7 DNA linear INV 07-OCT-2002
DEFINITION complete sequence.
ACCESSION AE014829 AE014185
VERSION   AE014829.1 GI:23494865
KEYWORDS
SOURCE    Plasmodium falciparum 3D7
ORGANISM  Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 250078)
AUTHORS   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
          Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T.,
          James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A.,
          Kyes S., Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J.,
          Angiuoli S., Pertea M., Allen J., Selengut J., Haft D.,
          Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H.,
          Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I.,
          Cummings L.M., Subramanian G.M., Mungall C., Venter J.C.,
          Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M. and
          Barrell B.

TITLE     Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL   Nature 419 (6906), 498-511 (2002)
PUBMED   12368864
REFERENCE 2 (bases 1 to 250078)
AUTHORS   Gardner M.J.
TITLE     Direct Submission
JOURNAL   Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
          Medical Center Dr., Rockville, MD 20850, USA

FEATURES             Location/Qualifiers
     source            1..250078
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="genomic DNA"
                     /isolate="3D7"
                     /db_xref="taxon:36329"
                     /chromosome="10"
     repeat_region     1..842
                     /rpt_type=tandem
     repeat_region     962..995
                     /rpt_type=tandem
     repeat_region     1165..1226
                     /rpt_type=tandem
     repeat_region     1165..1225
                     /rpt_type=tandem
     repeat_region     2096..2758
                     /rpt_type=tandem
     repeat_region     2261..2922
                     /rpt_type=tandem
     repeat_region     2990..3367
                     /rpt_type=tandem
     repeat_region     9425..9594
                     /rpt_type=tandem
     repeat_region     9426..9563
                     /rpt_type=tandem
     repeat_region     9556..9594
                     /rpt_type=tandem
     repeat_region     9840..9873
                     /rpt_type=tandem
     repeat_region     9880..9944
                     /rpt_type=tandem
     repeat_region     9886..9964
                     /rpt_type=tandem
     repeat_region     10025..10204
                     /rpt_type=tandem
     repeat_region     10079..10204
                     /rpt_type=tandem
     repeat_region     10080..10213
                     /rpt_type=tandem
     repeat_region     10225..10307
                     /rpt_type=tandem
     repeat_region     10301..10359
                     /rpt_type=tandem
     repeat_region     10383..10483
                     /rpt_type=tandem
     repeat_region     10396..10541
                     /rpt_type=tandem
     repeat_region     10693..10777
                     /rpt_type=tandem
     repeat_region     10719..10795
                     /rpt_type=tandem
     repeat_region     10870..11098
                     /rpt_type=tandem
     repeat_region     11196..11235
                     /rpt_type=tandem
     repeat_region     11251..11347
                     /rpt_type=tandem
     repeat_region     11383..12436
                     /rpt_type=tandem
     repeat_region     12702..17518
                     /rpt_type=tandem
     repeat_region     13365..18358
                     /rpt_type=tandem
     repeat_region     16920..19680
                     /rpt_type=tandem
     repeat_region     16935..20332
                     /rpt_type=tandem
     repeat_region     16935..19680
                     /rpt_type=tandem
     repeat_region     19528..19698
                     /rpt_type=tandem
     repeat_region     19703..19903
                     /rpt_type=tandem
     repeat_region     19987..20331
                     /rpt_type=tandem
     repeat_region     19987..20121
                     /rpt_type=tandem
     repeat_region     20045..20265
                     /rpt_type=tandem
     repeat_region     20175..20310
                     /rpt_type=tandem
     repeat_region     20402..20538
                     /rpt_type=tandem
     repeat_region     20446..20501
                     /rpt_type=tandem
     repeat_region     20455..20610
                     /rpt_type=tandem
     repeat_region     20605..20740
                     /rpt_type=tandem
     repeat_region     20622..20825
                     /rpt_type=tandem
     repeat_region     20786..20955
                     /rpt_type=tandem
     repeat_region     20935..21030
                     /rpt_type=tandem
     repeat_region     21037..21322
                     /rpt_type=tandem
     repeat_region     21269..21467
                     /rpt_type=tandem
     repeat_region     21529..21636
                     /rpt_type=tandem
     repeat_region     21590..21994
                     /rpt_type=tandem
     repeat_region     21625..21998
                     /rpt_type=tandem
     repeat_region     21632..22441
                     /rpt_type=tandem
```

repeat\_region 21632..22001  
/rpt\_type=tandem  
repeat\_region 21663..22441  
/rpt\_type=tandem  
repeat\_region 22727..22829  
/rpt\_type=tandem  
repeat\_region 22762..22842  
/rpt\_type=tandem  
repeat\_region 22875..23014  
/rpt\_type=tandem  
repeat\_region 22998..25249  
/rpt\_type=tandem  
repeat\_region 23670..23959  
/rpt\_type=tandem  
repeat\_region 23904..24125  
/rpt\_type=tandem  
repeat\_region 23985..24150  
/rpt\_type=tandem  
repeat\_region 24430..25018  
/rpt\_type=tandem  
repeat\_region 24527..24663  
/rpt\_type=tandem  
repeat\_region 24530..24751  
/rpt\_type=tandem  
repeat\_region 24617..24655  
/rpt\_type=tandem  
repeat\_region 25405..25446  
/rpt\_type=tandem  
repeat\_region 25464..25567  
/rpt\_type=tandem  
repeat\_region 25572..25734  
/rpt\_type=tandem  
repeat\_region 25763..26236  
/rpt\_type=tandem  
repeat\_region 26240..26353  
/rpt\_type=tandem  
repeat\_region 26382..26411  
/rpt\_type=tandem  
repeat\_region 26455..26494  
/rpt\_type=tandem  
repeat\_region 26541..26574  
/rpt\_type=tandem  
repeat\_region 26638..26753  
/rpt\_type=tandem  
repeat\_region 27339..27375  
/rpt\_type=tandem  
repeat\_region 28248..28336  
/rpt\_type=tandem  
Gene  
<28491..>36165  
/locus\_tag="PF10\_0001"  
/note="PF EMP1"  
join(<28491..33893,34843..>36165)  
/locus\_tag="PF10\_0001"  
join(28491..33893,34843..36165)  
/locus\_tag="PF10\_0001"  
/codon\_start=1  
/product="erythrocyte membrane protein 1 (PFEMP1)"  
/protein\_id="AA35201.1"  
/db\_xref="GI:23494866"  
/translation="MAAQSSGGGCGEEDKAYMFDRIKGVHDEVKNAANVFDY  
LKNLVTISTIGEETVAFTDPCELIKEKRDLLAARGDPCGKGHRFSKRGACDD  
NKEGNVRSGKGGKSAGACAPYRLSLCNKPNMNSKSLKAKNDLLVDVGYAAY  
EASLTPYHDQKLYGDSQICTVLARSFADIGDIVRGKDLFLGHNRKKLEERLEQ  
MFKNITHSNAKLSVLSSKIREYWMALNRDQVTKALTCDEENKLGNGAYPHATCSER  
NGGCSQAHEKCRCPWTSDBGKPNQVPTYFDYVQYLRLWFEWAEDFCRKKIYVDIVK

Query Match 2.8%; Score 30; DB 3; Length 250078;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCCTTTTGAAGAAAAAAGAAAAA 1079  
|||||  
DB 94906 TTTCCCTTTTGAAGAAAAAAGAAAAA 94935  
|||||

RESULT 7  
AE014818  
LOCUS  
DEFINITION  
Plasmodium falciparum 3D7 chromosome 14 section 3 of 13 of the complete sequence.  
ACCESSION  
AE014818 AE014187  
VERSION  
AE014818.1 GI:23497188  
KEYWORDS  
SOURCE  
Plasmodium falciparum 3D7  
ORGANISM  
Plasmodium falciparum 3D7  
REFERENCE  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
1 (bases 1 to 252632)  
AUTHORS  
Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McPadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.  
TITLE  
Genome sequence of the human malaria parasite Plasmodium falciparum  
JOURNAL  
Nature 419 (6906), 498-511 (2002)  
PUBMED  
12368864  
REFERENCE  
2 (bases 1 to 252632)  
AUTHORS  
Gardner,M.J.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (13-SEP-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
LOCATION/Qualifiers  
1 .252632  
/organism="Plasmodium falciparum 3D7"  
/mol\_type="genomic DNA"  
/isolate="3D7"  
/db\_xref="taxon:36329"  
/chromosome="14"  
76..110  
repeat\_region /rpt\_type=tandem  
140..166  
repeat\_region /rpt\_type=tandem  
828..857  
repeat\_region /rpt\_type=tandem  
1548..1607  
repeat\_region /rpt\_type=tandem  
1846..1913  
repeat\_region /rpt\_type=tandem  
2010..2087  
repeat\_region /rpt\_type=tandem  
2370..2401  
repeat\_region /rpt\_type=tandem  
2560..2613  
repeat\_region /rpt\_type=tandem  
2567..2627  
repeat\_region /rpt\_type=tandem  
2930..2964  
repeat\_region /rpt\_type=tandem  
3001..>4509  
/locus\_tag="PF14\_0124"  
join(<3001..3448,3827..>4509)  
/locus\_tag="PF14\_0124"  
join(3001..3448,3827..4509)  
/locus\_tag="PF14\_0124"  
/codon\_start=1  
/product="actin II"  
/protein\_id="AA36736.1"  
/db\_xref="GI:23497189"  
/translation="MSESAVALVVDNGSMVKSLAGDADPKCVFPSIVGRPKMNMIM  
IGMEQCYGVGDEAQNKRGIILTKYPIEHGIVTWNDDMEKIWHHTFYNELRVSPEHP  
VLLTEAPLNPKTNREKMTQIMFTFDVPMYVSIQAILSLYASGTTGTVLSDGSGVS  
HTVPIYEGYVLPAINRDMAGRDLTYHMKLFTERGHTFTTTTAREIVRDIKEXLCY







NEDFLLRKIKNAELIKKAKLKNHINIKRKTGPKPAAMIKVPPSKIIITIEPLPFYEE  
Y"

repeat\_region

8467. .8550

repeat\_region

/rpt\_type=tandem

repeat\_region

/rpt\_type=tandem

repeat\_region

8477. .8537

repeat\_region

/rpt\_type=tandem

repeat\_region

8723. .8844

repeat\_region

/rpt\_type=tandem

repeat\_region

8723. .8840

repeat\_region

/rpt\_type=tandem

repeat\_region

8782. .8844

repeat\_region

/rpt\_type=tandem

repeat\_region

9019. .9059

repeat\_region

/rpt\_type=tandem

repeat\_region

9019. .9058

repeat\_region

/rpt\_type=tandem

repeat\_region

9133. .9189

repeat\_region

/rpt\_type=tandem

repeat\_region

9221. .9254

repeat\_region

/rpt\_type=tandem

repeat\_region

9604. .9637

repeat\_region

/rpt\_type=tandem

repeat\_region

9668. .9723

repeat\_region

/rpt\_type=tandem

repeat\_region

10036. .10073

repeat\_region

/rpt\_type=tandem

repeat\_region

10445. .10472

repeat\_region

/rpt\_type=tandem

repeat\_region

10511. .10551

repeat\_region

/rpt\_type=tandem

gene

/rpt\_type=tandem

complement (<10565. .>11080)

locus\_tag="PF14\_0241"

mRNA

complement (<10565. .>11080)

CDS

locus\_tag="PF14\_0241"

complement (10565. .11080)

codon\_start=1

/product="basic transcription factor 3b, putative"

/protein\_id="AA036854.1"

/db\_xref="GI:23497309"

translation="MEKISPEILARAKLKERMGGNLRQIGKGARRKIKVHKNSI

SNKKINILIKIGASVFGVDIEICVYRTGDTFLFKPKLCAISLQSNYYIVTGKPN

HKIDINKLPEGLKGNKLDNMLEKIKNDPNIKNLKESGDTPKREEDQEANDVPD

LVNFEVSKE"

repeat\_region

11429. .11458

repeat\_region

/rpt\_type=tandem

repeat\_region

11514. .11584

repeat\_region

/rpt\_type=tandem

repeat\_region

11514. .11579

Query Match

2.7%; Score 29; DB 3; Length 250029;

Best Local Similarity

100.0%; Pred. No. 0.0008;

Matches 29; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCTTTTTCAGAAAAA

AAAAAAAAAAAA 1079

Db 42212 TTCTTTTTCAGAAAAA

AAAAAAAAAAAA 42184

RESULT 9

AL513496/c

LOCUS

AL513496

DEFINITION

Homo sapiens chromosome X clone RP3-359K12.

ACCESSION

AL513496

VERSION

AL513496.2 GI:13121498

KEYWORDS

HTG; HTGS PHASE2; HTGS\_CANCELLED.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

TITLE

JOURNAL

COMMENT

Lawlor,S.

Direct Submission

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgehire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

requests: clonerequest@sanger.ac.uk

On Feb 23, 2001 this sequence version replaced gi:12733864.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: dj359K12

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 118159 bases at least Q40

Consensus quality: 118179 bases at least Q30

Consensus quality: 118187 bases at least Q20

Insert size: 118189; sum-of-contigs

Insert size: 125554; 4.1% error; agarose-fp

Quality coverage: 9.82x in Q20 bases; sum-of-contigs Quality

coverage: 9.24x in Q20 bases; agarose-fp

-----

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

Source

Location/Qualifiers

1. .118189

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="X"

/clone="RP3-359K12"

/clone\_lib="RPCI-3"

1. .118189

/note="assembly\_fragment:03123"

misc\_feature

Query Match

Best Local Similarity

Matches 28; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TCCTTTTTCAGAAAAA

AAAAAAAAAAAA 1079

Db 54570 TCCTTTTTCAGAAAAA

AAAAAAAAAAAA 54543

RESULT 10

HS357K22/c

LOCUS

DEFINITION

Human DNA sequence from clone RP3-357K22 on chromosome Xq27.1-27.3,

complete sequence.

ACCESSION

AL022720

VERSION

AL022720.1 GI:3980360

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 145638)

Heath,P.

Direct Submission

Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Dec 8, 1998 this sequence version replaced gi:3646072.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>  
RP3-357K22 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: pCYPAC2

This sequence is the entire insert of clone RP3-357K22 The true left end of clone GS1-91018 is at 140544 in this sequence.

#### FEATURES

source	Location/Qualifiers	
	1. .145638	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="R2PD:RPCIP704K22357"	
	/db_xref="taxon:9606"	
	/chromosome="X"	
	/map="q27.1-27.3"	
	/clone="RP3-357K22"	
	/clone_lib="RPCI-3"	
repeat_region	2466. .2530	/note="HERVHF21 repeat: matches 27. .92 of consensus"
repeat_region	4030. .4077	/note="24 copies 2 mer ca 81% conserved"
repeat_region	5702. .5731	/note="15 copies 2 mer gt 86% conserved"
repeat_region	6735. .6788	/note="27 copies 2 mer gt 79% conserved"
repeat_region	7286. .7337	/note="26 copies 2 mer gt 90% conserved"
repeat_region	10302. .10349	/note="24 copies 2 mer ag 75% conserved"
repeat_region	12407. .14052	/note="MSTB-internal repeat: matches 1. .1651 of consensus"
misc_feature	18008. .18181	/note="match: STS: Em:L41319"
misc_feature	20159. .20189	/note="Tandem repeat. Forced join. From pac pcr's and restriction digest missing approx 20-30bp"
repeat_region	24708. .24833	/note="78K repeat: matches 2. .125 of consensus"
repeat_region	24875. .24942	/note="7SK repeat: matches 126. .193 of consensus"
repeat_region	30894. .30929	/note="18 copies 2 mer ac 100% conserved"
repeat_region	33712. .34079	/note="MLT12 repeat: matches 1. .374 of consensus"
repeat_region	35613. .35644	/note="16 copies 2 mer gt 84% conserved"
repeat_region	35854. .35907	/note="27 copies 2 mer ga 81% conserved"
repeat_region	37263. .37286	/note="12 copies 2 mer tt 100% conserved"
repeat_region	37488. .38795	/note="L1MBC repeat: matches 1448. .2414 of consensus"
repeat_region	43947. .44006	/note="L1MA7 repeat: matches 6228. .6285 of consensus"
repeat_region	45177. .45345	/note="MLT2 repeat: matches 1. .170 of consensus"
repeat_region	45346. .45933	/note="HERVJ repeat: matches 3179. .3767 of consensus"
repeat_region	45934. .46211	/note="MLT2 repeat: matches 175. .549 of consensus"
repeat_region	46226. .46715	/note="WER9 repeat: matches 1. .511 of consensus"
repeat_region	52697. .52722	/note="13 copies 2 mer tg 96% conserved"
repeat_region	52863. .53243	/note="THEIC repeat: matches 1. .371 of consensus"
repeat_region	53244. .54804	/note="THEIC-internal repeat: matches 1. .1580 of consensus"
repeat_region	54805. .55180	/note="THEIC repeat: matches 1. .371 of consensus"
repeat_region	56592. .56633	/note="21 copies 2 mer ac 100% conserved"
repeat_region	57083. .57207	/note="WER66A repeat: matches 393. .467 of consensus"
repeat_region	57351. .57723	/note="THEIC repeat: matches 1. .371 of consensus"
repeat_region	60523. .60878	/note="THEIA repeat: matches 1. .354 of consensus"
repeat_region	60879. .62452	/note="THEIA-internal repeat: matches 1. .1578 of consensus"
repeat_region	62455. .62809	/note="THEIA repeat: matches 1. .354 of consensus"
repeat_region	64319. .64568	/note="WER45C repeat: matches 692. .952 of consensus"
repeat_region	65051. .65182	/note="66 copies 2 mer aa 59% conserved"
repeat_region	66376. .66491	/note="58 copies 2 mer aa 59% conserved"
repeat_region	69569. .70134	/note="L1MCA repeat: matches 275. .844 of consensus"
repeat_region	70166. .70809	/note="L1MCh repeat: matches 625. .1294 of consensus"
repeat_region	70930. .71217	/note="144 copies 2 mer aa 56% conserved"
repeat_region	72738. .74181	/note="THE1B-INTERNAL repeat: matches 1. .1549 of consensus"
repeat_region	74182. .74621	/note="WER61B repeat: matches 1. .424 of consensus"
repeat_region	74628. .77796	/note="WER61-internal repeat: matches 1362. .4550 of consensus"
repeat_region	80415. .80777	/note="THEIC repeat: matches 1. .371 of consensus"
repeat_region	80778. .81041	/note="THEIC-internal repeat: matches 1. .264 of consensus"
repeat_region	81354. .82669	/note="THEIC-internal repeat: matches 264. .1580 of consensus"
repeat_region	82670. .83032	/note="THEIC repeat: matches 1. .371 of consensus"
repeat_region	88307. .88664	/note="L1PA2 repeat: matches 4473. .4830 of consensus"
repeat_region	88663. .89975	/note="L1PA2 repeat: matches 4833. .6145 of consensus"
repeat_region	90791. .90830	/note="20 copies 2 mer aa 82% conserved"
repeat_region	92362. .92389	/note="14 copies 2 mer tt 96% conserved"
misc_feature	complement(97390. .97547)	/note="match: STS: Em:L24645"
repeat_region	98507. .98567	/note="L1P8a repeat: matches -209. -.149 of consensus"
repeat_region	98530. .100503	/note="L1PA15-16 repeat: matches 566. .2456 of consensus"
repeat_region	106224. .106547	/note="L1MBC repeat: matches 1541. .1874 of consensus"

```

repeat_region 106552..106937
/note="L1MA8 repeat: matches 5894. .6284 of consensus"
repeat_region 107659..107726
/note="34 copies 2 mer aa 67% conserved"
repeat_region 108859..108916
/note="29 copies 2 mer ga 74% conserved"
repeat_region 114000..114043
/note="22 copies 2 mer ct 81% conserved"
repeat_region 115870..115899
/note="15 copies 2 mer tg 90% conserved"
repeat_region 116775..117108
/note="MT2A repeat: matches 2. .304 of consensus"
repeat_region 117092..117279
/note="MT2A repeat: matches 263. .453 of consensus"
repeat_region 121735..121786
/note="26 copies 2 mer at 80% conserved"
repeat_region 126415..126534
/note="60 copies 2 mer ta 76% conserved"
repeat_region 131714..131786
/note="78K repeat: matches 121. .193 of consensus"
repeat_region 131823..131949
/note="78K repeat: matches 1. .125 of consensus"
repeat_region 135220..135282
/note="L1PA3 repeat: matches 6084. .6146 of consensus"
repeat_region 144427..144948
/note="L1MA3 repeat: matches 5605. .6133 of consensus"
repeat_region 145088..145638
/note="L1PA5 repeat: matches 5596. .6145 of consensus"

```

## ORIGIN

```

Query Match      2.6%; Score 28; DB 9; Length 145638;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1052 TCCTTTTGAAGAAAAAAGAAAAA 1079
      |||
Db 112349 TCCTTTTGAAGAAAAAAGAAAAA 112322

```

## RESULT 11

```

AC135583
LOCUS      Rattus norvegicus clone CH230-219P9, *** SEQUENCING IN PROGRESS
DEFINITION Rattus norvegicus clone CH230-219P9, *** SEQUENCING IN PROGRESS
ACCESSION AC135583
VERSION    AC135583.2 GI:25139083
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

## REFERENCE

```

1 (bases 1 to 224697)
AUTHORS    Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

```

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naik,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemeloh,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanckoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puozzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R., Woodden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

```

Direct Submission
Unpublished
2 (bases 1 to 224697)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 224697)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

```

## COMMENT

On Nov 20, 2002 this sequence version replaced gi:24137403.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBVQ
Center clone name: CH230-219P9
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 209475 bases at least Q40
Consensus quality: 212691 bases at least Q30
Consensus quality: 214935 bases at least Q20
Estimated insert size: 218919; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length

```

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 223478: contig of 223478 bp in length  
 \* 223479 223578: gap of unknown length  
 \* 223579 224697: contig of 1119 bp in length.

#### FEATURES

source  
 1. 224697  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-219P9"

#### ORIGIN

Query Match 2.6%; Score 28; DB 2; Length 224697;  
 Best Local Similarity 100.0%; Pred. No. 0.0028;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCCTTTTGAAGAAAAA 1078  
 |||||  
 Db 74709 TTCCTTTTGAAGAAAAA 74736

#### RESULT 12

AC109654  
 LOCUS  
 Rattus norvegicus clone CH230-222B23, WORKING DRAFT SEQUENCE.  
 AC109654  
 AC109654.6 GI:30581398  
 HTG; HTGS PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

1 (bases 1 to 230038)

Muzny,D.,Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draber,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jacobs,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,

Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowig,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenshewa,L., Loulsegod,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindartne,M., Mahmoud,M., Mallory,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mathew,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Miloevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,B., Pfankoch,C.,  
 Plopper,F., Poidexter,A., Popovic,D., Primus,E., Pu,L.-L.,  
 Puazo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,  
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstock,G. and Gibbs,R.A.

#### Direct Submission

Unpublished  
 2 (bases 1 to 230038)  
 Worley,K.C.

#### Direct Submission

Submitted (07-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 230038)

#### Rat Genome Sequencing Consortium.

#### Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

#### COMMENT

On May 13, 2003 this sequence version replaced gi:22857421.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

#### Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRQO

Center clone name: CH230-222B23

Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 224292 bases at least Q40

Consensus quality: 225574 bases at least Q30

Consensus quality: 226622 bases at least Q20

Estimated insert size: 234782; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

#### \*\*\*\*\*

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 230038: contig of 230038 bp in length.



```
QY 1051 TTCCTTTTGAAGAAAAA 1078
Db 14669 TTCCTTTTGAAGAAAAA 14696

RESULT 14
CQ672612 207 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 17538 from Patent WO2070737.
DEFINITION CQ672612
ACCESSION CQ672612
VERSION CQ672612.1 GI:42169050
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 17538 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source 1..207
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.5%; Score 27; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.011; 0; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1053 CCTTTTGAAGAAAAA 1079
Db 154 CCTTTTGAAGAAAAA 180

RESULT 15
AX330437/c 446 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 946 from Patent WO0194629.
DEFINITION AX330437
ACCESSION AX330437
VERSION AX330437.1 GI:18103415
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 946 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1..446
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.5%; Score 27; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.011; 0; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1053 CCTTTTGAAGAAAAA 1079
Db 29 CCTTTTGAAGAAAAA 3

Search completed: September 12, 2005, 15:37:23
Job time : 4691.81 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 15:37:37 ; Search time 619.757 Seconds  
(without alignments)  
9663.988 Million cell updates/sec

Title: US-10-780-002-36  
Perfect score: 912  
Sequence: 1 atcagcgtattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 7351250 seqs, 3283620254 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :				Published Applications NA:*	
1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:				
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:				
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:				
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:				
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:				
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:				
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:				
8:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:				
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:				
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:				
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:				
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:				
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:				
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:				
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:				
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:				
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:				
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:				
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:				
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:				
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:				
22:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:				
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:				
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:				
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:				
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	912	100.0	912	21	US-10-780-002-36
2	432	47.4	483	10	US-09-770-961-854
3	22	2.4	2466	19	US-10-437-963-45846
4	21	2.3	1239	19	US-10-437-963-31574
5	20	2.2	201	20	US-10-719-993-14754
6	20	2.2	201	20	US-10-719-993-14862
7	20	2.2	8592	15	US-10-311-455-1955

Query Match	100.0%	Score 912;	DB 21;	Length 912;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 912;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATCAGCGTATTCTCAACAAAGCAAGACCTTAAACAAAGAAACCTCTACGAGACA	60	

RESULT 1  
US-10-780-002-36  
; Sequence 36, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleszig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; PRIOR FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 912  
; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-36

ALIGNMENTS

c	8	20	2.2	366710	20	US-10-719-993-6792	Sequence 6792, Ap
	9	19	2.1	319	19	US-10-767-701-24762	Sequence 24762, A
	10	19	2.1	345	19	US-10-437-963-98094	Sequence 98094, A
	11	19	2.1	387	16	US-10-116-712-82	Sequence 82, Appl
	12	19	2.1	418	20	US-10-425-115-77654	Sequence 77654, A
	13	19	2.1	432	19	US-10-767-795-5332	Sequence 5332, Ap
	14	19	2.1	440	16	US-10-116-712-643	Sequence 643, App
	15	19	2.1	462	19	US-10-437-963-58265	Sequence 58265, A
	16	19	2.1	624	9	US-09-902-941-1914	Sequence 1914, Ap
	17	19	2.1	624	9	US-09-849-626-1914	Sequence 1914, Ap
	18	19	2.1	624	14	US-10-017-754-1914	Sequence 1914, Ap
	19	19	2.1	624	16	US-10-113-872-1914	Sequence 1914, Ap
	20	19	2.1	624	17	US-10-283-017-1914	Sequence 1914, Ap
	21	19	2.1	636	9	US-09-736-457-1668	Sequence 1668, Ap
	22	19	2.1	636	9	US-09-902-941-1668	Sequence 1668, Ap
	23	19	2.1	636	9	US-09-849-626-1668	Sequence 1668, Ap
	24	19	2.1	636	14	US-10-017-754-1668	Sequence 1668, Ap
	25	19	2.1	636	16	US-10-113-872-1668	Sequence 1668, Ap
	26	19	2.1	636	17	US-10-283-017-1668	Sequence 1668, Ap
	27	19	2.1	1008	15	US-10-156-761-7148	Sequence 7148, Ap
	28	19	2.1	1310	17	US-10-264-049-218	Sequence 218, App
	29	19	2.1	1633	9	US-09-736-457-316	Sequence 316, App
	30	19	2.1	1633	9	US-09-736-457-788	Sequence 788, App
	31	19	2.1	1633	9	US-09-902-941-316	Sequence 316, App
	32	19	2.1	1633	9	US-09-902-941-788	Sequence 788, App
	33	19	2.1	1633	9	US-09-849-626-316	Sequence 316, App
	34	19	2.1	1633	9	US-09-849-626-788	Sequence 788, App
	35	19	2.1	1633	10	US-09-476-300-316	Sequence 316, App
	36	19	2.1	1633	14	US-10-017-754-316	Sequence 316, App
	37	19	2.1	1633	14	US-10-017-754-788	Sequence 788, App
	38	19	2.1	1633	16	US-10-113-872-316	Sequence 316, App
	39	19	2.1	1633	16	US-10-113-872-788	Sequence 788, App
	40	19	2.1	1633	17	US-10-283-017-316	Sequence 316, App
	41	19	2.1	1633	17	US-10-283-017-788	Sequence 788, App
	42	19	2.1	1633	17	US-10-172-118-1143	Sequence 1143, Ap
	43	19	2.1	1633	18	US-10-342-887-1143	Sequence 1143, Ap
	44	19	2.1	1633	21	US-10-848-755A-139	Sequence 139, App
	45	19	2.1	1679	20	US-10-817-483-17	Sequence 17, App1

Db 1 ATCAGCTATTCTCAACAGCAAGAACCTTTAAACCAAAAGAAAACGCTTACGGAGACA 60  
Qy 61 TATGGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTCGATCTG 120  
Db 61 TATGGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTCGATCTG 120  
Qy 121 GTACAAAGCTCAAGCCCTCTCTTGAATCAGCCGGCCACCGCTTACTGCTGCAACTGCG 180  
Db 121 GTACAAAGCTCAAGCCCTCTCTTGAATCAGCCGGCCACCGCTTACTGCTGCAACTGCG 180  
Qy 181 CCGCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 240  
Db 181 CCGCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 240  
Qy 241 ACCGTTGATCGAAACCCCTCAAAATCTCTTCAGAGAACGAAGAGGTAATTCGTGGTGA 300  
Db 241 ACCGTTGATCGAAACCCCTCAAAATCTCTTCAGAGAACGAAGAGGTAATTCGTGGTGA 300  
Qy 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 360  
Db 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 360  
Qy 361 TCTTGTTCTCCTCAAGCCCTCTTGGAGATTTGGGAGATTTGTGAGTTTTCATCTCATGA 420  
Db 361 TCTTGTTCTCCTCAAGCCCTCTTGGAGATTTGGGAGATTTGTGAGTTTTCATCTCATGA 420  
Qy 421 CAAATGATATGAGATGAGCTTCGAGGTTTGGGAGATTTGTGAGTTTTCATCTCATGA 480  
Db 421 CAAATGATATGAGATGAGCTTCGAGGTTTGGGAGATTTGTGAGTTTTCATCTCATGA 480  
Qy 481 AAATGGAGCATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 540  
Db 481 AAATGGAGCATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 540  
Qy 541 AAATGTTCCATAGAGGATTTACGAGCTGSCAAAATGTTGCATAGCAAGGCTCATTTT 600  
Db 541 AAATGTTCCATAGAGGATTTACGAGCTGSCAAAATGTTGCATAGCAAGGCTCATTTT 600  
Qy 601 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGAGATATGTTTCGGTCAACG 660  
Db 601 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGAGATATGTTTCGGTCAACG 660  
Qy 661 AGTTTACGTAAGTAGTAGAGACAAAGCCATCCCTCGCATTTTCATTCGTTGATGAT 720  
Db 661 AGTTTACGTAAGTAGTAGAGACAAAGCCATCCCTCGCATTTTCATTCGTTGATGAT 720  
Qy 721 TGAATATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCATGCTGATGCT 780  
Db 721 TGAATATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCATGCTGATGCT 780  
Qy 781 CTCCTCAACCCCAAAACCTCTTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
Db 781 CTCCTCAACCCCAAAACCTCTTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
Qy 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTTACTTAATAAAACCCCTTTTTCGG 900  
Db 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTTACTTAATAAAACCCCTTTTTCGG 900  
Qy 901 GCAACTTTCATC 912  
Db 901 GCAACTTTCATC 912

RESULT 2

US-09-770-961-854  
; Sequence 854, Application US/09770961  
; Publication No. US20030115639A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kriker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffmann, Neil  
; APPLICANT: Hurlban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE OF INVENTION: thaliana  
; FILE REFERENCE: 2026 (PARA-015PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,961  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,466  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 854  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-961-854

Query Match 47.4%; Score 432; DB 10; Length 483;

Best Local Similarity 99.8%; Pred. No. 8.4e-233; Indels 0; Gaps 0;  
Matches 482; Conservative 0; Mismatches 1;

Qy 26 GAACTTAAACCAAAAGAAAACGCTTACGGAGACATATGGAGAGAAACATCACTTCGTG 85  
Db 1 GAACTTAAACCAAAAGAAAACGCTTACGGAGACATATGGAGAGAAACATCACTTCGTG 60  
Qy 86 TTAGTTCAACAGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGCCCTCTTTGAA 145  
Db 61 TTAGTTCAACAGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGCCCTCTTTGAA 120  
Qy 146 TCAGCGCCACCGCTTACTGCTGAACTCCGCGCTCCGGGATCGACCCACGACCA 205  
Db 121 TCAGCGCCACCGCTTACTGCTGAACTCCGCGCTCCGGGATCGACCCACGACCA 180  
Qy 206 ATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAAACCCCTCAAATCT 265  
Db 181 ATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAAACCCCTCAAATCT 240  
Qy 266 CTTCCAGAGAACGAAGAGGTAATTCGTGGATTTCAGCTTCGGAGGATCAACATCGCT 325  
Db 241 CTTTCAGAGAACGAAGAGGTAATTCGTGGATTTCAGCTTCGGAGGATCAACATCGCT 300  
Qy 326 CTCGCGCCGACATATTTCCGCGCAAGATTAAAGTTCTTGTTCTTCCCAAGCCCTCTTG 385  
Db 301 CTCGCGCCGACATATTTCCGCGCAAGATTAAAGTTCTTGTTCTTCCCAAGCCCTCTTG 360  
Qy 386 CCCGACCAACCCACGCTTCCTCAGCTTCGGACGAAGTATATGGAGATCCCTGGAGGT 445  
Db 361 CCCGACCAACCCACGCTTCCTCAGCTTCGGACGAAGTATATGGAGATCCCTGGAGGT 420  
Qy 446 TTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGGGACGATGATTTATTGAAG 505  
Db 421 TTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGGGACGATGATTTATTGAAG 480  
Qy 506 ATG 508  
Db 481 ATG 483

RESULT 3

US-10-437-963-45846  
; Sequence 45846, Application US/10437963



Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 45846  
LENGTH: 2466  
TYPE: DNA  
ORGANISM: Oryza sativa  
NAME/KEY: unsure  
LOCATION: (1)..(2466)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48771C.1  
US-10-437-963-45846

Query Match 2.4%; Score 22; DB 19; Length 2466;  
Best Local Similarity 100.0%; Pred. No. 0.95;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 CTGCGCGCCTCCGGATCGACC 197  
DB 382 CTGCGCGCCTCCGGATCGACC 403

RESULT 4  
US-10-437-963-31574  
Sequence 31574, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 31574  
LENGTH: 1239  
TYPE: DNA  
ORGANISM: Oryza sativa  
NAME/KEY: unsure  
LOCATION: (1)..(1239)  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_35868C.1  
US-10-437-963-31574

Query Match 2.3%; Score 21; DB 19; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 611 CTATCAAGAGAAAGTTT 631  
DB 622 CTATCAAGAGAAAGTTT 642

RESULT 5  
US-10-719-993-14754  
Sequence 14754, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
CURRENT FILING DATE: 2003-11-24  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14754  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-14754

Query Match 2.2%; Score 20; DB 20; Length 201;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 ATTGAAGATGGACCAAAAT 518  
DB 127 ATTGAAGATGGACCAAAAT 146

RESULT 6  
US-10-719-993-14862  
Sequence 14862, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
CURRENT FILING DATE: 2003-11-24  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14862  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-14862

Query Match 2.2%; Score 20; DB 20; Length 201;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 ATTGAAGATGGACCAAAAT 518  
DB 13 ATTGAAGATGGACCAAAAT 32

RESULT 7  
US-10-311-455-1955/c  
Sequence 1955, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIRPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30

```
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1955
; LENGTH: 8592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1955
```

```
Query Match          2.2%; Score 20; DB 15; Length 8592;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 781 CTCCAAAACCCCAAACTCT 800
Db 2474 CTCCAAAACCCCAAACTCT 2455
|||||
```

## RESULT 8

```
US-10-719-993-6792
; Sequence 6792, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6792
; LENGTH: 366710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(366710)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6792
```

```
Query Match          2.2%; Score 20; DB 20; Length 366710;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 499 ATTGAAGATGGGACCAAAAT 518
Db 184736 ATTGAAGATGGGACCAAAAT 184755
|||||
```

## RESULT 9

```
US-10-767-701-24762/c
; Sequence 24762, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24762
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30941244
US-10-767-701-24762
```

```
Query Match          2.1%; Score 19; DB 19; Length 319;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 169 TGTGAACTGCGCGCTCC 187
Db 145 TGTGAACTGCGCGCTCC 127
|||||
```

## RESULT 10

```
US-10-437-963-98094
; Sequence 98094, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 98094
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96032C.1
US-10-437-963-98094
```

```
Query Match          2.1%; Score 19; DB 19; Length 345;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 771 TGGTGATGCTCTCCAAACC 789
Db 234 TGGTGATGCTCTCCAAACC 252
|||||
```

## RESULT 11

```
US-10-116-712-82/c
; Sequence 82, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 320_336, 345
; OTHER INFORMATION: n = A,T,C or G
US-10-116-712-82
```

```
Query Match          2.1%; Score 19; DB 16; Length 387;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 137 CCTTTTCCGGGCAACTTT 119

RESULT 12  
US-10-425-115-77654/c  
; Sequence 77654, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 77654  
; LENGTH: 418  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_170847C.1  
US-10-425-115-77654

Query Match 2.1%; Score 19; DB 20; Length 418;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 AAAACAAACCCCTTTTCC 898  
|||||  
Db 133 AAAACAAACCCCTTTTCC 115

RESULT 13  
US-10-767-795-5332  
; Sequence 5332, Application US/10767795  
; Publication No. US20040181830A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10/767,795  
; CURRENT FILING DATE: 2004-01-30  
; NUMBER OF SEQ ID NOS: 117596  
; SEQ ID NO 5332  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3145-010-Q1-K1-E3  
US-10-767-795-5332

Query Match 2.1%; Score 19; DB 19; Length 432;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 AAAACAAACCCCTTTTCC 898  
|||||  
Db 1 AAAACAAACCCCTTTTCC 19

RESULT 14  
US-10-116-712-643/c  
; Sequence 643, Application US/10116712  
; Publication No. US20030194764A1

; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Switzer, Ann  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.568  
; CURRENT APPLICATION NUMBER: US/10/116,712  
; CURRENT FILING DATE: 2002-04-07  
; NUMBER OF SEQ ID NOS: 670  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 643  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-116-712-643

Query Match 2.1%; Score 19; DB 16; Length 440;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 137 CCTTTTCCGGGCAACTTT 119

RESULT 15  
US-10-437-963-58265/c  
; Sequence 58265, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 58265  
; LENGTH: 462  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_5999C.1  
US-10-437-963-58265

Query Match 2.1%; Score 19; DB 19; Length 462;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 AATCCAGGCCGTTGAACCC 223  
|||||  
Db 217 AATCCAGGCCGTTGAACCC 199

Search completed: September 12, 2005, 20:50:01  
Job time : 623.757 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:35:08 ; Search time 173.147 Seconds  
(without alignments)  
8618.590 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcacgctattctcaacaa.....ttttcgggcaactttcatc 912

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	2.4	601	US-09-949-016-183626	Sequence 183626,
2	22	2.4	45983	US-09-949-016-17010	Sequence 17010, A
3	19	2.1	393	US-09-540-236-856	Sequence 856, App
4	19	2.1	636	US-09-702-705-1668	Sequence 1668, App
5	19	2.1	636	US-09-736-457-1668	Sequence 1668, App
6	19	2.1	636	US-09-614-124B-1668	Sequence 1668, App
7	19	2.1	636	US-09-671-325-1668	Sequence 1668, App
8	19	2.1	636	US-09-658-824-1668	Sequence 1668, App
9	19	2.1	1633	US-09-702-705-316	Sequence 316, App
10	19	2.1	1633	US-09-702-705-788	Sequence 788, App
11	19	2.1	1633	US-09-736-457-316	Sequence 316, App
12	19	2.1	1633	US-09-736-457-788	Sequence 788, App
13	19	2.1	1633	US-09-614-124B-316	Sequence 316, App
14	19	2.1	1633	US-09-614-124B-788	Sequence 788, App
15	19	2.1	1633	US-09-671-325-316	Sequence 316, App
16	19	2.1	1633	US-09-671-325-788	Sequence 788, App
17	19	2.1	1633	US-09-589-184-316	Sequence 316, App
18	19	2.1	1633	US-09-589-184-788	Sequence 788, App
19	19	2.1	1633	US-09-658-824-316	Sequence 316, App
20	19	2.1	1633	US-09-658-824-788	Sequence 788, App
21	19	2.1	1635	US-09-949-016-4625	Sequence 4625, App
22	19	2.1	9544	US-09-949-016-16367	Sequence 16367, A
23	18	2.0	333	US-09-248-796A-2805	Sequence 2805, App
24	18	2.0	366	US-09-248-796A-11450	Sequence 11450, A
25	18	2.0	601	US-09-949-016-20860	Sequence 20860, A
26	18	2.0	601	US-09-949-016-110207	Sequence 110207,
27	18	2.0	601	US-09-949-016-198581	Sequence 198581,

ALIGNMENTS

RESULT 1

US-09-949-016-183626  
; Sequence 183626, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 183626

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-183626

Query Match 2.4%; Score 22; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 AGGAACATCACTTCGTTGTTAG 89

DB 512 AGGAACATCACTTCGTTGTTAG 533

RESULT 2

US-09-949-016-17010/c

; Sequence 17010, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

```
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17010
/ LENGTH: 45983
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..- (45983)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17010

Query Match      2.4%; Score 22; DB 4; Length 45983;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      68 AGGAACATCACTTCGTGTAG 89
Db      34394 AGGAACATCACTTCGTGTAG 34373

RESULT 3
US-09-540-236-856/c
/ Sequence 856, Application US/09540236
/ Patent No. 6673910
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2005-001
/ CURRENT APPLICATION NUMBER: US/09/540,236
/ CURRENT FILING DATE: 2000-04-04
/ NUMBER OF SEQ ID NOS: 3840
/ SEQ ID NO 856
/ LENGTH: 393
/ TYPE: DNA
/ ORGANISM: M.catarrhalis
US-09-540-236-856

Query Match      2.1%; Score 19; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      785 AAACCCCAAAACTCTTTG 803
Db      272 AAACCCCAAAACTCTTTG 254

RESULT 4
US-09-702-705-1668/c
/ Sequence 1668, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C14
/ CURRENT APPLICATION NUMBER: US/09/702,705
/ CURRENT FILING DATE: 2000-10-30
/ NUMBER OF SEQ ID NOS: 1833
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1668
/ LENGTH: 636
/ TYPE: DNA
```

```
/ ORGANISM: Homo sapiens
US-09-702-705-1668

Query Match      2.1%; Score 19; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      890 CCTTTTCCGGGCAACTTT 908
Db      562 CCTTTTCCGGGCAACTTT 544

RESULT 5
US-09-736-457-1668/c
/ Sequence 1668, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1668
/ LENGTH: 636
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-736-457-1668

Query Match      2.1%; Score 19; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      890 CCTTTTCCGGGCAACTTT 908
Db      562 CCTTTTCCGGGCAACTTT 544

RESULT 6
US-09-614-124B-1668/c
/ Sequence 1668, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1668
/ LENGTH: 636
/ TYPE: DNA
/ ORGANISM: Homo sapiens
```

US-09-614-124B-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 7

US-09-671-325-1668/c  
; Sequence 1668, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C12  
; CURRENT APPLICATION NUMBER: US/09/671,325  
; CURRENT FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 1825  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1668  
; LENGTH: 636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-671-325-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 8

US-09-658-824-1668/c  
; Sequence 1668, Application US/09658824  
; Patent No. 6746846  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C11  
; CURRENT APPLICATION NUMBER: US/09/658,824  
; CURRENT FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 1788  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1668  
; LENGTH: 636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-658-824-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 9

US-09-702-705-316/c  
; Sequence 316, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-702-705-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 10

US-09-702-705-788/c  
; Sequence 788, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-702-705-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 11

US-09-736-457-316/c  
; Sequence 316, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Lijun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 12

US-09-736-457-788/c  
; Sequence 788, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Lijun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 13

US-09-614-124B-316/c  
; Sequence 316, Application US/09614124B  
; Patent No. 6630574  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C9  
; CURRENT APPLICATION NUMBER: US/09/614,124B  
; CURRENT FILING DATE: 2001-07-11  
; NUMBER OF SEQ ID NOS: 1668  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-614-124B-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 14

US-09-614-124B-788/c  
; Sequence 788, Application US/09614124B  
; Patent No. 6630574  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C9  
; CURRENT APPLICATION NUMBER: US/09/614,124B  
; CURRENT FILING DATE: 2001-07-11  
; NUMBER OF SEQ ID NOS: 1668  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-614-124B-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 08:41:37 ; Search time 2908.23 Seconds  
(without alignments)  
11936.677 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcaactattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hc:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gsa1:.\*  
9: gb\_gsa2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	718	78.7	865	CNS09Z6U	BX833804 Arabidops
2	628	68.9	667	AV822206	AV822206 Arabidops
3	624	68.4	625	AV785311	AV785311 AV785311
4	554	60.7	641	CB258477	CB258477 05-E01273
5	435	47.7	537	AV783452	AV783452 AV783452
6	421	46.2	421	BP613009	BP613009 BP613009
7	397	43.5	572	AV822640	AV822640 AV822640
8	374	41.0	441	AV824147	AV824147 AV824147
9	366	40.1	366	AV554161	AV554161 AV554161
10	341	37.4	390	AV820476	AV820476 AV820476
11	289	31.7	480	AA04853	AA04853 28251 Lam
12	279	30.6	335	AV545429	AV545429 AV545429
13	263	28.8	382	BP646852	BP646852 BP646852
14	259	28.4	361	AI993246	AI993246 701495839
15	247	27.1	308	AV559909	AV559909 AV559909
16	232	25.4	405	BP642613	BP642613 BP642613
17	232	25.4	425	BP642196	BP642196 BP642196
18	217	23.8	217	AV782897	AV782897 AV782897
19	203	22.3	351	BP604502	BP604502 BP604502
20	167	18.3	426	AV790585	AV790585 AV790585
21	142	15.6	296	CNS00PYH	AL085127 Arabidops
22	136	14.9	444	BX292336	BX292336 Arabidops
23	99	10.9	488	BX292222	BX292222 Arabidops
24	87	9.5	159	AI996681	AI996681 701667925

C	25	85	9.3	581	9	BX292335	BX292335 Arabidops
	26	50	5.5	177	7	R84081	R84081 16040 Lambd
	27	33	3.6	556	8	BZ507980	BZ507980 BONA086TR
	28	32	3.5	138	9	BX891283	BX891283 Arabidops
C	29	25	2.7	411	5	BP652964	BP652964 BP652964
	30	23	2.5	730	8	BZ237627	BZ237627 CH230-333
C	31	22	2.4	433	2	BE328824	BE328824 hv97601.X
	32	22	2.4	698	9	CL567704	CL567704 OB_Ba003
C	33	22	2.4	725	8	BZ050121	BZ050121 jnr44c05
	34	22	2.4	1083	9	CL200471	CL200471 ZMMBRC023
	35	21	2.3	418	6	CB765657	CB765657 AMGNNUC:N
	36	21	2.3	539	1	AL921928	AL921928 AL921928
C	37	21	2.3	557	8	BH822165	BH822165 BACP17-M
	38	21	2.3	605	2	BF099324	BF099324 601751709
C	39	21	2.3	681	7	CK414051	CK414051 AUF_IpGii
	40	21	2.3	691	7	CK413120	CK413120 AUF_IpGii
	41	21	2.3	760	2	BF241640	BF241640 601878974
	42	21	2.3	822	5	BQ432919	BQ432919 AGENCOURT
C	43	21	2.3	855	9	CL851022	CL851022 OR_CBA007
	44	21	2.3	889	6	CA474406	CA474406 AGENCOURT
C	45	21	2.3	1176	9	AG340901	AG340901 MUB_muscu

#### ALIGNMENTS

CNS09Z6U 865 bp mRNA linear HTC 06-FBB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTSL782C11 of Silique of strain col-0 of Arabidopsis thaliana  
(thale cress).

ACCESSION BX833804

VERSION BX833804.1 GI:42455725

KEYWORDS HTC; GSLT cDNA.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1 (bases 1 to 865)

AUTHORS Castell, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,

Menard, G., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 865)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen)

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,

URV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.

Location/Qualifiers

1..865

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strains="Col-0"

/db\_xref="taxon:3702"

```
/clone="GSLTSL178ZC11"
/tissue_type="Silique"
/plasmid="pCMVSPORT_6"
1..865
/gene="At5g10300"

Query Match      78.7%; Score 718; DB 3; Length 865;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 838; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 75 ATCACTTCGTGTAGTTCACACGCTTATCATGAGCGCTTGATCTGGTACAGCTCAAGCTCAAGC 134
DB 1 ATCACTTCGTGTAGTTCACACGCTTATCATGAGCGCTTGATCTGGTACAGCTCAAGCTCAAGC 60

QY 135 CCCTCCTTTGAATCAGCGCGCCACCGCTTACTGTGCGAACTCGCCGCCCTCCGGGATCG 194
DB 61 CCCTCCTTTGAATCAGCGCGCCACCGCTTACTGTGCGAACTCGCCGCCCTCCGGGATCG 120

QY 195 ACCCAGCAATTCACGCGGTTGAAACCGTTCGACGAATATCTCCAAACCGTTGATCGAAA 254
DB 121 ACCCAGCAATTCACGCGGTTGAAACCGTTCGACGAATATCTCCAAACCGTTGATCGAAA 180

QY 255 CCCTCAATCTCTCCAGAGACGAGGTAATTCGTGGTTCAGTTCAGCTTCGGAGGCA 314
DB 181 CCCTCAATCTCTCCAGAGACGAGGTAATTCGTGGTTCAGTTCAGCTTCGGAGGCA 240

QY 315 TCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTCTTGTTCTCTCA 374
DB 241 TCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTCTTGTTCTCTCA 300

QY 375 ACGCTCTTTCGCCGACACAA - CCAACGTCCTTCTCAGCTTCTGGACAAGTATATGGAG 433
DB 301 ACGCTCTTTCGCCGACACAAACCCACGTCCTTCTCAGCTTCTGGACAAGTATATGGAG 360

QY 434 ATGCTGGAGGTTGGAGATTGTGAGTTTTCATCTCATGAACCAAGAAATGGGACGATG 493
DB 361 ATGCTGGAGGTTGGAGATTGTGAGTTTTCATCTCATGAACCAAGAAATGGGACGATG 420

QY 494 AGTTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTTACCAAAATGTCCCATTA 553
DB 421 AGTTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTTACCAAAATGTCCCATTA 480

QY 554 GAGGATTACAGTGGCAAAAATGTTGCATAGCGAAGGTCTATTTTTCACAGAGATCTA 613
DB 481 GAGGATTACAGTGGCAAAAATGTTGCATAGCGAAGGTCTATTTTTCACAGAGATCTA 540

QY 614 TCAAGAAAGAAAGTTTAGCGGAAGGATATGTTTCGGTGCACGAGTTTACGTTAATG 673
DB 541 TCAAGAAAGAAAGTTTAGCGGAAGGATATGTTTCGGTGCACGAGTTTACGTTAATG 600

QY 674 AGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATTAATTTCAAC 733
DB 601 AGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATTAATTTCAAC 660

QY 734 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGTATGCTCTCCTCAAAACCCCAA 793
DB 661 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGTATGCTCTCCTCAAAACCCCAA 720

QY 794 AAACCTTTTGACTCTCTCTGCTATTCGCCAGGATATATGTAATATCTTAAGTCGT 853
DB 721 AAACCTTTTGACTCTCTCTGCTATTCGCCAGGATATATGTAATATCTTAAGTCGT 780

QY 854 TTCTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 912
DB 781 TTCTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 839

RESULT 2
AV822206 667 bp mRNA linear EST 01-APR-2002
LOCUS AV822206 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-02-P20 5',
DEFINITION mRNA sequence.

ACCESSION AV822206
VERSION AV822206.1 GI:19864240
EST. Arabidopsis thaliana (thale cress)
KEYWORDS Arabidopsis thaliana
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 667)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.
FEATURES
Location/Qualifiers
1..667
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-02-P20"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated (1,2,5,10,24 hr)"
ORIGIN
Query Match 68.9%; Score 628; DB 1; Length 667;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AAGAACCTTACCAAAAGAAACGCTCTACGAGACATATCGAGGGAACATCACATTCG 83
DB 25 AAGAACCTTACCAAAAGAAACGCTCTACGAGACATATCGAGGGAACATCACATTCG 84

QY 84 TGTTAGTTCAACACGCTTATCATGGAGCTCGATCTGTGTAAGCTCAAGCCCTCCTTG 143
DB 85 TGTTAGTTCAACACGCTTATCATGGAGCTCGATCTGTGTAAGCTCAAGCCCTCCTTG 144

QY 144 AATCAGCGCGCACCGCGTTACTGTGTGAACTCGCCCTCGGGATCGACCCACGAC 203
DB 145 AATCAGCGCGCACCGCGTTACTGTGTGAACTCGCCCTCGGGATCGACCCACGAC 204

QY 204 CAATCCAGGCGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAAACCTCAAT 263
DB 205 CAATCCAGGCGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAAACCTCAAT 264

QY 264 CTCCTCCAGAGAACGAGAGGTAACTCTGGTTGGATTCAGCTTCGGAGGCAATCAACATCG 323
DB 265 CTCCTCCAGAGAACGAGAGGTAACTCTGGTTGGATTCAGCTTCGGAGGCAATCAACATCG 324

QY 324 CTCCTCCGCGGACATATTTCCGGGGAAGATTAAGGTTCTTGTTCTCAACGCTTCT 383
DB 325 CTCCTCCGCGGACATATTTCCGGGGAAGATTAAGGTTCTTGTTCTCAACGCTTCT 384

QY 384 TGCCCGACACAAACCCACGTCCTTCTACGTTCTGGACAAGTATATGAGATGCTCGAG 443
DB 385 TGCCCGACACAAACCCACGTCCTTCTACGTTCTGGACAAGTATATGAGATGCTCGAG 444

QY 444 GTTTCGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATGAGTTTATGA 503
```

Db 445 GTTTGGGAGATTCTGAGTTTTCATCTCATGAAACAAGAAATGGACGATGAGTTTATTGA 504  
 QY 504 AGATGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATGTCTCCATAGAGGATTAGC 563  
 Db 505 AGATGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATGTCTCCATAGAGGATTAGC 564  
 QY 564 AGCTGGCAAAATGTTCATAGCAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAG 623  
 Db 565 AGCTGGCAAAATGTTCATAGCAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAG 624  
 QY 624 AAAAGTTTACGAGGAGGATGATGGTTC 651  
 Db 625 AAAAGTTTACGAGGAGGATGATGGTTC 652

RESULT 3  
 AV785311/c  
 LOCUS  
 DEFINITION AV785311 RAPL6 Arabidopsis thaliana cDNA clone RAPL06-14-F12 3', mRNA linear EST 28-MAR-2002  
 mRNA sequence.  
 ACCESSION AV785311 GI:19804101  
 VERSION  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 625)  
 REFERENCE  
 AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
 Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.  
 and Shinozaki,K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished (2002)  
 CONTACT: Motoaki Seki  
 RIKEN Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). This clone is in a  
 modified pBluescript vector as a SstI/XhoI insert. Please visit our  
 web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for  
 further details.

FEATURES  
 source  
 1..625  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAPL06-14-F12"  
 /dev\_stage="plants at various developmental stages from  
 germination to mature seeds"  
 /lab\_host="DH10B"  
 /clone\_lib="RAPL6"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to  
 dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
 hr) treatments"

ORIGIN  
 Query Match 68.4%; Score 624; DB 1; Length 625;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 279 AAGAGGTAATTCGTTGGATTTCAGCTTCGAGGCATCAACATCGCTCTCGCGCGGACA 338  
 Db 625 AAGAGGTAATTCGTTGGATTTCAGCTTCGAGGCATCAACATCGCTCTCGCGCGGACA 566  
 QY 339 TATTTCCGGCGAAGATTAAAGGTTCTTGTTCTCTCAACGCCCTTTCTTGCCCGCACAAACC 398

Db 565 TATTTCCGGCGAAGATTAAAGGTTCTTGTTCTCTCAACGCCCTTTCTTGCCCGCACAAACC 506  
 QY 399 AGTGCCCTTCTCACGTTCTGGCAAGTATATGAGATGCTCGAGGTTTGGGAGATTGTG 458  
 Db 505 AGTGCCCTTCTCACGTTCTGGCAAGTATATGAGATGCTCGAGGTTTGGGAGATTGTG 446  
 QY 459 AGTTTTCATCTCATGAAAACAAGAAATGGACGATGAGTTTATTTGAAGATGGACCAAAAT 518  
 Db 445 AGTTTTCATCTCATGAAAACAAGAAATGGACGATGAGTTTATTTGAAGATGGACCAAAAT 386  
 QY 519 TCATGAAGCAGCTCTTTACCAAAATTTGTCATAGAGGATTACGAGCTGGCAAAATGT 578  
 Db 385 TCATGAAGCAGCTCTTTACCAAAATTTGTCATAGAGGATTACGAGCTGGCAAAATGT 326  
 QY 579 TGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 638  
 Db 325 TGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 266  
 QY 639 AAGGATATGGTTTCGATGCAAGGATTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCT 698  
 Db 265 AAGGATATGGTTTCGATGCAAGGATTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCT 206  
 QY 699 GCGATTTTCATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATG 758  
 Db 205 GCGATTTTCATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATG 146  
 QY 759 GCGAGATCACATGATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTA 818  
 Db 145 GCGAGATCACATGATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTA 86  
 QY 819 TTGCCACCGATTATATGTAATAATCTTAAGTCCGTTTTTCTATCTTTTCTCATCGTTACTAA 878  
 Db 85 TTGCCACCGATTATATGTAATAATCTTAAGTCCGTTTTTCTATCTTTTCTCATCGTTACTAA 26  
 QY 879 TAAACAAACCCCTTTTTCGCGGC 902  
 Db 25 TAAACAAACCCCTTTTTCGCGGC 2

RESULT 4  
 CB258477  
 LOCUS  
 DEFINITION CB258477 641 bp mRNA linear EST 06-NOV-2003  
 clone MPI2p771J012Q 5-PRIME, mRNA sequence.  
 ACCESSION CB258477  
 VERSION CB258477.1 GI:32883250  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 641)  
 REFERENCE  
 AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,  
 Mitchell-Olds,T. and Weisshaar,B.  
 Large-scale identification and analysis of genome-wide  
 single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
 Genome Res. 13 (6), 1250-1257 (2003)  
 22683290  
 MEDLINE  
 PUBMED 12799357  
 COMMENT Contact: Weisshaar B  
 ADIS DNA core facility at MPIZ  
 Max-Planck-Institute for Plant Breeding Research  
 Carl-von-Linne Weg 10, 50829 Koeln, Germany  
 Fax: 00492215062851  
 Email: weisshaar@mpiz-koeln.mpg.de  
 Insert Length: 641 Std Error: 0.00  
 Plate: 2 row: J column: 01  
 Seq primer: T7R; CTAATCGACTCACTATAGGA.  
 Location/Qualifiers  
 1..641  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"



```

QY 616 AAGAAAGAAAAGCTTTAGCGAGGAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 675
|
Db 297 AAGAAAGAAAAGCTTTAGCGAGGAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 238
|
QY 676 TAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCAACGT 735
|
Db 237 TAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCAACGT 178
|
QY 736 CTCGAAAGTCTACGAGATCGATGGCGAGATCACATGGTGATGCTCTCCAAACCCCAAAA 795
|
Db 177 CTCGAAAGTCTACGAGATCGATGGCGAGATCACATGGTGATGCTCTCCAAACCCCAAAA 118
|
QY 796 ACTCTTTGACTCTCTCTCTGCTAFTGCCACCGATTATATGTAATAATCTTAAGTCGGTTT 855
|
Db 117 ACTCTTTGACTCTCTCTCTGCTAFTGCCACCGATTATATGTAATAATCTTAAGTCGGTTT 58
|
QY 856 TACTTTTTTCTCATCGTTACTATAAACAACAAACCCCTTTTCCGGCAACTTTTCATC 912
|
Db 57 TACTTTTTTCTCATCGTTACTATAAACAACAAACCCCTTTTCCGGCAACTTTTCATC 1

RESULT 6
BP613009/c
LOCUS BP613009 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-91-K04 3',
DEFINITION mRNA sequence.
ACCESSION BP613009.1 GI:49264191
VERSION EST.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 421)
AUTHORS Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
TITLE Functional annotation of a full-length Arabidopsis cDNA collection
MEDLINE Science 296 (5565), 141-145 (2002)
PUBMED 21932900
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
reversed clone; please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
source
location/Qualifiers
1..421
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL16-91-K04"
/lab_host="DH10B"
/clone_lib="RAFL16"
/note="Site 1: BamHI; Site 2: SalI; dark-grown"

ORIGIN
Query Match 46.2%; Score 421; DB 5; Length 421;
Best Local Similarity 100.0%; Pred. No. 6:3e-222;
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 TGAGTTTATTGAAGATGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCCA 551
|
Db 421 TGAGTTTATTGAAGATGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCCA 362
|
QY 552 TAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGGTCATTTTTCACAGAGGATC 611
|

```

```

Db 361 TAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGGTCATTTTTCACAGAGGATC 302
|
QY 612 TATCAAGAAAGAAAAGTTTACGAGGAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 671
|
Db 301 TATCAAGAAAGAAAAGTTTACGAGGAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 242
|
QY 672 TGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCA 731
|
Db 241 TGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCA 182
|
QY 732 AGCTCTCGAAAGTCTACGAGATCGATGGCGAGATCACATGGTGATGCTCTCCAAACCC 791
|
Db 181 AGCTCTCGAAAGTCTACGAGATCGATGGCGAGATCACATGGTGATGCTCTCCAAACCC 122
|
QY 792 AAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAAGTC 851
|
Db 121 AAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAAGTC 62
|
QY 852 GTTTTACTTTTTTCTCATCGTTACTATAAACAACAAACCCCTTTTCCGGGCAACTTTTCAT 911
|
Db 61 GTTTTACTTTTTTCTCATCGTTACTATAAACAACAAACCCCTTTTCCGGGCAACTTTTCAT 2

QY 912 C 912
|
Db 1 C 1

RESULT 7
AV822640
LOCUS AV822640 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-09-N03 5',
DEFINITION mRNA sequence.
ACCESSION AV822640
VERSION AV822640.1 GI:19864692
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 572)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.
location/Qualifiers
1..572
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-09-N03"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration-treated(1,2,5,10,24 hr)"

ORIGIN
Query Match 43.5%; Score 397; DB 1; Length 572;

```

Best Local Similarity 99.6%; Pred. No. 1.4e-208;		Matches 567; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
QY		1 ATCAGCCTATTCTCAACAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 60	
Db		2 ATCAGCCTATTCTCAACAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 61	
QY		61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 120	
Db		62 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 121	
QY		121 GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGTGTGAACTGCG 180	
Db		122 GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGTGTGAACTGCG 181	
QY		181 CGCTCCGGGATCGACCCAGCAACCAATCCAGCCGTTGAAACCGTCGACGAATATCCAA 240	
Db		182 CGCTCCGGGATCGACCCAGCAACCAATCCAGCCGTTGAAACCGTCGACGAATATCCCA 241	
QY		241 ACCGTTGATCGAAACCTCTTCCAGAGAACGAAGAGGTAAATTCGTGTTGGATT 300	
Db		242 ACCGTTGATCGAAACCTCTTCCAGAGAACGAAGAGGTAAATTCGTGTTGGATT 301	
QY		301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTCGGCGAAGATTAAAGT 360	
Db		302 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTCGGCGAAGATTAAAGT 361	
QY		361 TCTTGTTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA 420	
Db		362 TCTTGTTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA 421	
QY		421 CAAAGT 425	
Db		422 CAAAGT 426	
RESULT 9		AV554161	
LOCUS		AV554161	
DEFINITION		Arabidopsis thaliana (thale cress)	
ACCESSION		AV554161	
VERSION		AV554161.1	
KEYWORDS		EST.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
TITLE		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL		1 (bases 1 to 441)	
COMMENT		Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,	
		Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,	
		Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.	
		and Shinozaki, K.	
		Large scale analysis of Arabidopsis full-length cDNA (2002b)	
		Unpublished (2002)	
		Contact: Motoaki Seki	
		Plant Functional Genomics Research Group	
		RIKEN Genomic Sciences Center	
		3-1-1 Toyodai, Tsukuba, Ibaraki 305-0074, Japan	
		Tel: 81-298-36-4359	
		Fax: 81-298-36-9060	
		Email: msek@rtc.riken.go.jp	
		An Arabidopsis full-length cDNA library was constructed essentially	
		as reported previously (Seki et al., 1998). This clone is in a	
modified pBluescript vector as a SstI/XhoI insert. Please visit our		web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for	
further details.		Location/Qualifiers	
FEATURES		1..441	
source		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clones="RAF106-14-F12"	
		/dev_stages="plants at various developmental stages from	
		germination to mature seeds"	
		/lab_host="DH10B"	
		/clone_lib="RAPL6"	
		/note="Site 1: SstI; Site 2: XhoI; subjected to	
		dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24	
		hr) treatments"	
ORIGIN		Query Match 41.0%; Score 374; DB 1; Length 441;	
		Best Local Similarity 99.8%; Pred. No. 8.2e-196;	
		Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY		1 ATCAGCCTATTCTCAACAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 60	
Db		2 ATCAGCCTATTCTCAACAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 61	
QY		61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 120	
Db		62 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 121	
QY		121 GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGTGTGAACTGCG 180	
Db		122 GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGTGTGAACTGCG 181	
QY		181 CGCTCCGGGATCGACCCAGCAACCAATCCAGCCGTTGAAACCGTCGACGAATATCCAA 240	
Db		182 CGCTCCGGGATCGACCCAGCAACCAATCCAGCCGTTGAAACCGTCGACGAATATCCCA 241	
QY		241 ACCGTTGATCGAAACCTCTTCCAGAGAACGAAGAGGTAAATTCGTGTTGGATT 300	
Db		242 ACCGTTGATCGAAACCTCTTCCAGAGAACGAAGAGGTAAATTCGTGTTGGATT 301	
QY		301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTCGGCGAAGATTAAAGT 360	
Db		302 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTCGGCGAAGATTAAAGT 361	
QY		361 TCTTGTTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA 420	
Db		362 TCTTGTTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA 421	
QY		421 CAAAGT 425	
Db		422 CAAAGT 426	
RESULT 9		AV554161	
LOCUS		AV554161	
DEFINITION		Arabidopsis thaliana roots Columbia Arabidopsis thaliana	
ACCESSION		CDNA clone R28203R 5', mRNA sequence.	
VERSION		AV554161	
KEYWORDS		EST.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
TITLE		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL		1 (bases 1 to 366)	
COMMENT		Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.	
		A large scale analysis of cDNA in Arabidopsis thaliana: Generation	
		of 12,028 non-redundant expressed sequence tags from normalized and	
		size-selected cDNA libraries	



JOURNAL DNA Res. 7 (3), 175-180 (2000)  
MEDLINE 20363093  
PUBMED 10907847  
COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
source  
1..366  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="R282h03R"  
/tissue type="roots"  
/clone\_lib="Arabidopsis thaliana roots Columbia"  
/notes="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 40.1%; Score 366; DB 1; Length 366;  
Best Local Similarity 100.0%; Pred. No. 2.3e-191;  
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 72 AACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCTGATCTGGTACAAGCTCA 131  
DB 1 AACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCTGATCTGGTACAAGCTCA 60  
QY 132 AGCCCTCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCAACTCGCGCTCCGGGA 191  
DB 61 AGCCCTCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCAACTCGCGCTCCGGGA 120  
QY 192 TCGACCCAGCAACATCCAGCGGTTGAAACCGTCCAGCAATCTCCAAACCGTTGATCG 251  
DB 121 TCGACCCAGCAACATCCAGCGGTTGAAACCGTCCAGCAATCTCCAAACCGTTGATCG 180  
QY 252 AAACCTCTCAATCTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTTCAGCTTCGGAG 311  
DB 181 AAACCTCTCAATCTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTTCAGCTTCGGAG 240  
QY 312 GCATCAACATCGCTTCGCGCGGACATATTTCCGCGGAGAGATTAAGGTTCTTGTTCC 371  
DB 241 GCATCAACATCGCTTCGCGCGGACATATTTCCGCGGAGATTAAGGTTCTTGTTCC 300  
QY 372 TCAACGCTCTTTCGCGGACACACCGTCTCTCAGCTTCGGAACAGTATATGG 431  
DB 301 TCAACGCTCTTTCGCGGACACACCGTCTCTCAGCTTCGGAACAGTATATGG 360  
QY 432 AGATGC 437  
DB 361 AGATGC 366

RESULT 10  
AV820476/c  
LOCUS AV820476 390 bp mRNA linear EST 01-APR-2002  
DEFINITION AV820476 RAF111 Arabidopsis thaliana cDNA clone RAF111-10-III 3', mRNA sequence.  
ACCESSION AV820476  
VERSION AV820476.1 GI:19862446  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 390)  
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, S., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinozaki, K., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)

JOURNAL Unpublished (2002)  
COMMENT Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msek@gtc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambdaBam-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further details.

FEATURES  
Location/Qualifiers  
1..390  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAF111-10-III"  
/dev stage="plants at various developmental stages from germination to mature seeds"  
/lab host="DH10B"  
/clone\_lib="RAF111"  
/note="Site 1: BamHI; Site 2: SalI; subjected to various treatments (dehydration, cold, high salt, ABA, heat and UV). Dark-grown plants"

ORIGIN  
Query Match 37.4%; Score 341; DB 1; Length 390;  
Best Local Similarity 100.0%; Pred. No. 1.8e-177;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 572 AAAATGTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTT 631  
DB 341 AAAATGTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTT 282  
QY 632 AGCAGGAGGAGATATGTTTCGTCGCAACGAGTTTACGTAATGAGTAGTGAACAAAGCC 691  
DB 281 AGCAGGAGGAGATATGTTTCGTCGCAACGAGTTTACGTAATGAGTAGTGAACAAAGCC 222  
QY 692 ATCCCTCTCGATTTTCATTCGTTGGATGATGATTAATTTCAACGCTTCGAAAGTCTACGAG 751  
DB 221 ATCCCTCTCGATTTTCATTCGTTGGATGATGATTAATTTCAACGCTTCGAAAGTCTACGAG 162  
QY 752 ATCGATGCGGAGATCACAATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTC 811  
DB 161 ATCGATGCGGAGATCACAATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTC 102  
QY 812 TCTGCTATTGCCACCGATTATATGTAATCTTAAGTCCGTTTACATTTTCTCATCG 871  
DB 101 TCTGCTATTGCCACCGATTATATGTAATCTTAAGTCCGTTTACATTTTCTCATCG 42  
QY 872 TTACTATAAAACAAACCCCTTTTTCGCGCAACTTTTCATC 912  
DB 41 TTACTATAAAACAAACCCCTTTTTCGCGCAACTTTTCATC 1

RESULT 11  
AA404853  
LOCUS 28251 Lambda-PRL2 Arabidopsis thaliana cDNA clone 143M1XP 3', mRNA sequence.  
DEFINITION AA404853  
ACCESSION AA404853  
VERSION AA404853.1 GI:2062871  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 480)

**AUTHORS** Newman,T., deBruijn,F.J., Green,P., Keestra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,  
Retzel,E. and Somerville,C.  
**TITLE** Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
**JOURNAL** Plant Physiol. 106, 1241-1255 (1994)  
**MEDLINE** 95148729  
**PUBMED** 7846151  
**COMMENT** Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@bm.cl.msu.edu  
Seq primer: M13 -21 dye primer.

**FEATURES**

source  
1..480  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="143M11XP"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: Lambda Zip-Lox; Site 1: Sal; Site 2: Not;  
Lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques). The vector is BRL's lambda Zip-Lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA."

**ORIGIN**

Query Match 31.7%; Score 289; DB 1; Length 480;  
Best Local Similarity 99.5%; Pred. No. 1.4e-148;  
Matches 389; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 512 CCATAATTCATGAGCAGCTCTTTACCAAAATGTCCTCATAGAGGATTACGAGCTGGCA 571  
DB 90 CCATAATTCATGAGCAGCTCTTTACCAAAATGTCCTCATAGAGGATTACGAGCTGGCA 149  
  
QY 572 AAAATGTTGCATAGGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTT 631  
DB 150 AAAATGTTGCATAGGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTT 209  
  
QY 632 AGCGAGGAAGGATATGGTTCGGTGCACGAGTTTACGTAATGAGTAGTGAAGACAAAGCC 691  
DB 210 AGCGAGGAAGGATATGGTTCGGTGCACGAGTTTACGTAATGAGTAGTGAAGACAAAGCC 269  
  
QY 692 ATCCCTCGGATTCATTCGTTGGATGATTGTAATTTCAACGCTCCGAAAGCTTACGAG 751  
DB 270 ATCCCTCGGATTCATTCGTTGGATGATTGTAATTTCAACGCTCCGAAAGCTTACGAG 329  
  
QY 752 ATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTTTTGACTCTCTC 811  
DB 330 ATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTTTTGACTCTCTC 389  
  
QY 812 TCTGCTATTGCCACCGATTATATGTAATATCTTAAAGTCGGTTTACTTTTTTCTCATCG 871  
DB 390 TCTGCTATTGCCACCGATTATATGTAATATCTTAAAGTCGGTTTACTTTTTTCTCATCG 449  
  
QY 872 TTACTAATAAACAACACCCCTTTTCCGGGC 902  
DB 450 TTACTAATAAACAACACCCCTTTTCCGGGC 480

RESULT 12  
AV545429/c  
LOCUS AV545429 335 bp mRNA linear EST 20-FEB-2004

**DEFINITION** AV545429 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
cDNA clone R282h03F 3', mRNA sequence.  
**ACCESSION** AV545429  
**VERSION** AV545429.1 GI:8716843  
**KEYWORDS** EST.  
**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
**REFERENCE** 1 (bases 1 to 335)  
**AUTHORS** Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
**TITLE** A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
**JOURNAL** DNA Res. 7 (3), 175-180 (2000)  
**MEDLINE** 20363093  
**PUBMED** 10907847  
**COMMENT** Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

**FEATURES**

source  
1..335  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clones="R282h03F"  
/tissue\_type="roots"  
/clone\_lib="Arabidopsis thaliana roots Columbia"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

**ORIGIN**

Query Match 30.6%; Score 279; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 5e-143;  
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 634 CGAGGAAGGATATGGTTCGGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCAT 693  
DB 314 CGAGGAAGGATATGGTTCGGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCAT 255  
  
QY 694 CCCCTGCCATTCATTCGTTGGATGATTGTAATTTCAACGCTCGAAGTCTACGAGAT 753  
DB 254 CCCCTGCCATTCATTCGTTGGATGATTGTAATTTCAACGCTCGAAGTCTACGAGAT 195  
  
QY 754 CGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTTTTGACTCTCTCTC 813  
DB 194 CGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTTTTGACTCTCTCTC 135  
  
QY 814 TGCTATTGCCACCGATTATATGTAATATCTTAAAGTCGGTTTACTTTTTTCTCATCGTT 873  
DB 134 TGCTATTGCCACCGATTATATGTAATATCTTAAAGTCGGTTTACTTTTTTCTCATCGTT 75  
  
QY 874 ACTAATAAACAACACCCCTTTTCCGGGCAACTTTTCATC 912  
DB 74 ACTAATAAACAACACCCCTTTTCCGGGCAACTTTTCATC 36

RESULT 13  
BP646852/c  
LOCUS BP646852

**DEFINITION** BP646852 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-73-109 3',  
mRNA sequence.  
**ACCESSION** BP646852  
**VERSION** BP646852.1 GI:49298322  
**KEYWORDS** EST.

**ORGANISM** Arabidopsis thaliana (thale cress)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 382)  
 Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.  
 Functional annotation of a full-length Arabidopsis cDNA collection  
 Science 296 (5565), 141-145 (2002)

# REFERENCE

## AUTHORS

# TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 reversed clone; Please visit our web site  
 (http://pfweb.gsc.riken.go.jp/) for further details.

Location/Qualifiers

## FEATURES

## source

1. .382  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAF19-73-I09"  
 /tissue type="mixture of silique and flower"  
 /lab\_host="DH10B"  
 /clone lib="RAF19"  
 /notes="Site\_1: BamHI; Site\_2: SalI; Subtraction Library"

## ORIGIN

Query Match 28.8%; Score 263; DB 5; Length 382;  
 Best Local Similarity 99.7%; Pred. No. 3.9e-134;  
 Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 587 CAAGGTCATTTTTCACAGAGATCTATCAAGAAAGAAAGTTAGCGAGGAGATAT 646  
 Db 344 CAAGGTCATTTTTCACAGAGATCTATCAAGAAAGAAAGTTAGCGAGGAGATAT 285  
 QY 647 GGTTCGGTCAACGAGTTACGTAAAGTAGTAGTGAAGCAAGCCATCCCTGCCATTTC 706  
 Db 284 GGTTCGGTCAACGAGTTACGTAAAGTAGTAGTGAAGCAAGCCATCCCTGCCATTTC 225  
 QY 707 ATTCTGTGGATGATTCATTAATTTCAACGCTCGAAGTCTACGAGATCGATGCGGAGAT 766  
 Db 224 ATTCTGTGGATGATTCATTAATTTCAACGCTCGAAGTCTACGAGATCGATGCGGAGAT 165  
 QY 767 CACATGGTGATGCTCTCCAAACCCCAAACTTTTGACTCTCTCTGCTATTGCCACC 826  
 Db 164 CACATGGTGATGCTCTCCAAACCCCAAACTTTTGACTCTCTCTGCTATTGCCACC 105  
 QY 827 GATTATATGTAATCTTAAGTCGCGTTTACTTTTCTCATCGTTACTATAAACA 886  
 Db 104 GATTATATGTAATCTTAAGTCGCGTTTACTTTTCTCATCGTTACTATAAACA 45  
 QY 887 ACCCTTTTTCGGG 900  
 Db 44 ACCCTTTTTCGGG 31

## RESULT 14

## AI993246

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

AI993246 361 bp mRNA linear EST 08-SEP-1999  
 701495839 A. thaliana, Ohio State clone set Arabidopsis thaliana  
 cDNA clone 701495839, mRNA sequence.  
 AI993246  
 AI993246.1 GI:5840151  
 EST.  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

# REFERENCE

## AUTHORS

1 (bases 1 to 361)  
 Chen, J., Hillyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Millman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrigo, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.  
 Arabidopsis thaliana Gene Expression MicroArray  
 Unpublished (1999)  
 Contact: David Smoller, Ph.D.  
 Genome Systems, Inc., a wholly owned subsidiary of Incyte  
 Pharmaceuticals, Inc.  
 4633 World Parkway Circle, St. Louis, MO 63134, USA  
 Tel: 877-577-2733  
 Fax: 314-427-3324  
 Email: service@genomesystems.com.

# TITLE

## JOURNAL

## COMMENT

Contact: David Smoller, Ph.D.  
 Genome Systems, Inc., a wholly owned subsidiary of Incyte  
 Pharmaceuticals, Inc.  
 4633 World Parkway Circle, St. Louis, MO 63134, USA  
 Tel: 877-577-2733  
 Fax: 314-427-3324  
 Email: service@genomesystems.com.

## FEATURES

## source

1. .361  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="701495839"  
 /clone lib="A. thaliana, Ohio State clone set"  
 /note="cDNA library was made from selected clones from the  
 Arabidopsis thaliana Ohio State clone set."

## ORIGIN

Query Match 28.4%; Score 259; DB 1; Length 361;  
 Best Local Similarity 99.4%; Pred. No. 6.5e-132;  
 Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 14 CAACCAACCAAGAACCTTAACCAAAAGAAAACGCTACGAGACATATGGAGAGAAA 73  
 Db 1 CAACCAACCAAGAACCTTAACCAAAAGAAAACGCTACGAGACATATGGAGAGAAA 60  
 QY 74 CATCACTTCGTGTAGTTACAAAGCTTATCATGGAGCTGATCTGGTACAGTCAAG 133  
 Db 61 CATCACTTCGTGTAGTTACAAAGCTTATCATGGAGCTGATCTGGTACAGTCAAG 120  
 QY 134 CCCCTCTTGAATCAGCCGCCACCGCGTTACTGCTGCGAACTCGCCGCTCCGGGATC 193  
 Db 121 CCCCTCTTGAATCAGCCGCCACCGCGTTACTGCTGCGAACTCGCCGCTCCGGGATC 180  
 QY 194 GACCACCAACCAATCCAGCGCGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA 253  
 Db 181 GACCACCAACCAATCCAGCGCGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA 240  
 QY 254 ACCCTCAAAATCTCTTCCAGAGAACGAGAGGTAATTTCTGGTGGATTTCAGCTTCGAGGC 313  
 Db 241 ACCCTCAAAATCTCTTCCAGAGAACGAGAGGTAATTTCTGGTGGATTTCAGCTTCGAGGC 300  
 QY 314 ATCAACATCGCTCTCGCCGCCGACATATTTCCGGGGAAGATTAGGTTCTTGTGTCTTC 373  
 Db 301 ATCAACATCGCTCTCGCCGCCGACATATTTCCGGGGAAGATTAGGTTCTTGTGTCTTC 360  
 QY 374 A 374  
 Db 361 A 361

## RESULT 15

## AV559909/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

AV559909 308 bp mRNA linear EST 23-FEB-2004  
 AV559909 Arabidopsis thaliana green silicles Columbia Arabidopsis  
 thaliana cDNA clone SQ125f05F 3', mRNA sequence.  
 AV559909  
 AV559909.1 GI:8731335  
 EST.  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

resoids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 308)
AUTHORS	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL	DNA Res. 7 (3), 175-180 (2000)
MEDLINE	20363093
PUBMED	10907847
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES	Location/Qualifiers
source	1..308
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/ecotype="Columbia"
	/db_xref="taxon:3702"
	/clone="SQ125f05F"
	/tissue_type="green siliques"
	/clone_lib="Arabidopsis thaliana green siliques Columbia"
	/note="Vector: pbluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
ORIGIN	
Query Match	27.1%; Score 247; DB 1; Length 308;
Best Local Similarity	100.0%; Pred. No. 3e-125;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	666 ACGTAATCAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTCATA 725
Db	308 ACGTAATCAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTCATA 249
Qy	726 ATTTCAACGTCGCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGATGCTCTCCA 785
Db	248 ATTTCAACGTCGCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGATGCTCTCCA 189
Qy	786 AACCCCAAAAACCTCTTTGACCTCTCTCTGCTATATGCCACCGATTATATGTAATAATCTT 845
Db	188 AACCCCAAAAACCTCTTTGACCTCTCTCTGCTATATGCCACCGATTATATGTAATAATCTT 129
Qy	846 AAGTCGGTTTTACTTTTTTCTCATCGTTACTTAATAAAACAACCCCTTTTTCGGGGCAAC 905
Db	128 AAGTCGGTTTTACTTTTTTCTCATCGTTACTTAATAAAACAACCCCTTTTTCGGGGCAAC 69
Qy	906 TTTCATC 912
Db	68 TTTCATC 62

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 06:32:02 ; Search time 499.287 Seconds  
(without alignments)  
10813.032 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagcgtattctcaacaa.....ttttcgggcaacttctatc 912

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	912	100.0	984	8	ADA26534	Arabidops
2	861	94.4	1056	3	AAC47895	Arabidops
3	844	92.5	909	3	AAC47800	Arabidops
4	432	47.4	483	10	ADE82083	Arabidops
5	20	2.2	8592	6	ABL33982	Human imm
6	19	2.1	371	4	AAL80037	Human pol
7	19	2.1	387	10	ADD49350	Human lun
8	19	2.1	393	12	ADL03170	DNA encod
9	19	2.1	432	13	ADR64551	Cotton cd
10	19	2.1	440	10	ADD49911	Human lun
11	19	2.1	624	6	ABK39793	DNA encod
12	19	2.1	624	8	ACA12122	Human His
13	19	2.1	624	8	ACA03308	Lung canc
14	19	2.1	624	10	ADH47433	Human lun
15	19	2.1	624	13	ADJ211352	Human lun
16	19	2.1	636	6	ABK39629	cDNA enco
17	19	2.1	636	8	ACAL1958	Human lun
18	19	2.1	636	8	ACA03144	Lung canc
19	19	2.1	636	10	ADH47187	Human lun
20	19	2.1	636	13	ADJ211106	Human lun

C 21	19	2.1	1310	6	ABQ54338	Human ova
C 22	19	2.1	1633	5	AAP68398	Human lun
C 23	19	2.1	1633	5	AAP68850	Human lun
C 24	19	2.1	1633	6	ABK38761	cDNA enco
C 25	19	2.1	1633	6	ABK38309	cDNA enco
C 26	19	2.1	1633	8	ACA10638	Human lun
C 27	19	2.1	1633	8	ACA11090	Human lun
C 28	19	2.1	1633	8	ACA02276	Lung canc
C 29	19	2.1	1633	8	ABX99589	Lung canc
C 30	19	2.1	1633	10	ADH45835	Human lun
C 31	19	2.1	1633	10	ADH46307	Human lun
C 32	19	2.1	1633	12	ADE72372	Human lun
C 33	19	2.1	1633	13	ADR25282	Breast ca
C 34	19	2.1	1633	13	ADJ19754	Human lun
C 35	19	2.1	1633	13	ADJ20226	Human lun
C 36	19	2.1	2179	5	AAS87251	DNA encod
C 37	19	2.1	3266	4	AAB34750	Human col
C 38	19	2.1	3300	4	ABL03694	Drosophila
C 39	19	2.1	3521	8	ABX76282	Lung canc
C 40	19	2.1	3565	12	ADN41807	Novel hum
C 41	19	2.1	9325	13	ADR84244	Aspergill
C 42	19	2.1	32186	5	AAS34422	Human DNA
C 43	19	2.1	38855	12	ADN41808	Novel hum
C 44	19	2.1	63248	11	ACN45100	Mouse gen
C 45	18	2.0	283	6	ABL74493	Corn tags

#### ALIGNMENTS

RESULT 1  
ADA26534  
ID ADA26534 standard; DNA; 984 BP.  
XX  
AC ADA26534;

DT 20-NOV-2003 (first entry)

DE Arabidopsis thaliana AtSBL5 gene homologous to tobacco SABP2 gene.

XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.

OS Arabidopsis thaliana.

PN WO2003016551-A2.

XX 27-FEB-2003.

PF 16-AUG-2002; 2002WO-US026312.

XX 16-AUG-2001; 2001US-0312863P.

PR (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX Klessig DF, Kumar D;

DR WPI; 2003-278578/27.

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.

XX Claim 65; Page; 98pp; English.

PS The invention relates to a novel isolated 1079 bp salicylic acid-binding  
protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
protein, sequences that specifically hybridize with the nucleic acid, the  
complement of the nucleic acid or a natural allelic variant of the  
sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
conferring increased resistance to disease in plants, or for screening  
modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L5 which  
CC has homology to the gene (ADA26496) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).  
XX  
SQ Sequence 984 BP; 291 A; 230 C; 200 G; 263 T; 0 U; 0 Other;

Query Match		100.0%;	Score 912;	DB 8;	Length 984;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 912;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCAGCGTATTCTCAAAACAGCAAGAACCTTTAACCACAAAGAAACCGTCTACGGAGACA	60		
Db	1	ATCAGCGTATTCTCAAAACAGCAAGAACCTTTAACCACAAAGAAACCGTCTACGGAGACA	60		
Qy	61	TATGAGAGGAAACATCACTTCGTGTAGTTTACAAACGTTATCATGAGAGCTGGATCTG	120		
Db	61	TATGAGAGGAAACATCACTTCGTGTAGTTTACAAACGTTATCATGAGAGCTGGATCTG	120		
Qy	121	GTACAAGCTCAAGCCCTCTTGAATCAGCGGCCACCGCTTACTGCTGCGAACTCGC	180		
Db	121	GTACAAGCTCAAGCCCTCTTGAATCAGCGGCCACCGCTTACTGCTGCGAACTCGC	180		
Qy	181	CGCTTCCGGGATCGACCCACGACCAATCCAGCGCGTTGAAACCGTCGACGAATATCCAA	240		
Db	181	CGCTTCCGGGATCGACCCACGACCAATCCAGCGCGTTGAAACCGTCGACGAATATCCAA	240		
Qy	241	ACCGTTGATCGAAACCTTCAAAATCTCTCCAGAGAACGAAAGGTAATCTCTGGTTGGATT	300		
Db	241	ACCGTTGATCGAAACCTTCAAAATCTCTCCAGAGAACGAAAGGTAATCTCTGGTTGGATT	300		
Qy	301	CAGCTTCGGAGCATCAATCGCTCTCGCCGCGACATATTTCCGGCGAGATTAAAGT	360		
Db	301	CAGCTTCGGAGCATCAATCGCTCTCGCCGCGACATATTTCCGGCGAGATTAAAGT	360		
Qy	361	TCTTGTGTTCTCAACGCCCTTCTGCGCGACACCAACCGTGCCTTCTCAGTTCTGGA	420		
Db	361	TCTTGTGTTCTCAACGCCCTTCTGCGCGACACCAACCGTGCCTTCTCAGTTCTGGA	420		
Qy	421	CAAGTATATGAGATGCGCTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG	480		
Db	421	CAAGTATATGAGATGCGCTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG	480		
Qy	481	AAATGGGACGATGAGTTTATTCAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA	540		
Db	481	AAATGGGACGATGAGTTTATTCAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA	540		
Qy	541	AAATTTGTCCTATAGAGGATTACGAGTTCGCAAAAATGTTGCATAGGCAAGGTCATTTT	600		
Db	541	AAATTTGTCCTATAGAGGATTACGAGTTCGCAAAAATGTTGCATAGGCAAGGTCATTTT	600		
Qy	601	CACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAAGATATGTTTCGGTCAACG	660		
Db	601	CACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAAGATATGTTTCGGTCAACG	660		
Qy	661	AGTTTACGTAATGAGTAGTAGAGCAAGCAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT	720		
Db	661	AGTTTACGTAATGAGTAGTAGAGCAAGCAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT	720		
Qy	721	TGATAATTTCAACGCTCTCGAAGTCTTACGAGATCGATGGCGGAGATCACATGGTGATGCT	780		
Db	721	TGATAATTTCAACGCTCTCGAAGTCTTACGAGATCGATGGCGGAGATCACATGGTGATGCT	780		
Qy	781	CTCCAAACCCCAAAACCTTTTGAATCTCTCTCTGCTATTGCGACCGATTATATGTAATA	840		
Db	781	CTCCAAACCCCAAAACCTTTTGAATCTCTCTCTGCTATTGCGACCGATTATATGTAATA	840		
Qy	841	ATCTTAAGTCCGTTTTTACTTTTTTCTCATCGTTACTTAATAAAACCAACCCCTTTTCCGG	900		

Db	841	ATCTTAAGTCGGTTTACTTTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTTCGG	900	
Qy	901	GCAACTTTTCATC 912		
Db	901	GCAACTTTTCATC 912		
RESULT 2				
AAC47895				
ID	AAC47895 standard; DNA; 1056 BP.			
XX	AC	AAC47895;		
XX				
DT	18-OCT-2000 (first entry)			
XX				
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55510.			
XX				
KW	Hybridisation assay; genetic mapping; gene expression control;			
KW	protein identification; signal transduction pathway; metabolic pathway;			
KW	promoter; termination sequence; ss.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-00301439.			
XX				
PR	25-FEB-1999;	99US-0121825P.		
PR	09-MAR-1999;	99US-0123180P.		
PR	09-MAR-1999;	99US-0123548P.		
PR	23-MAR-1999;	99US-0125788P.		
PR	25-MAR-1999;	99US-0126264P.		
PR	29-MAR-1999;	99US-0126785P.		
PR	01-APR-1999;	99US-0127462P.		
PR	06-APR-1999;	99US-0128234P.		
PR	08-APR-1999;	99US-0128714P.		
PR	16-APR-1999;	99US-0129845P.		
PR	19-APR-1999;	99US-0130077P.		
PR	21-APR-1999;	99US-0130449P.		
PR	23-APR-1999;	99US-0130510P.		
PR	23-APR-1999;	99US-0130891P.		
PR	28-APR-1999;	99US-0131449P.		
PR	30-APR-1999;	99US-0132048P.		
PR	04-MAY-1999;	99US-0132407P.		
PR	05-MAY-1999;	99US-0132484P.		
PR	05-MAY-1999;	99US-0132485P.		
PR	06-MAY-1999;	99US-0132486P.		
PR	06-MAY-1999;	99US-0132487P.		
PR	07-MAY-1999;	99US-0132863P.		
PR	11-MAY-1999;	99US-0134256P.		
PR	14-MAY-1999;	99US-0134218P.		
PR	14-MAY-1999;	99US-0134219P.		
PR	14-MAY-1999;	99US-0134221P.		
PR	14-MAY-1999;	99US-0134370P.		
PR	18-MAY-1999;	99US-0134768P.		
PR	19-MAY-1999;	99US-0134941P.		
PR	20-MAY-1999;	99US-0135124P.		
PR	21-MAY-1999;	99US-0135353P.		
PR	24-MAY-1999;	99US-0135629P.		
PR	25-MAY-1999;	99US-0136021P.		
PR	27-MAY-1999;	99US-0136392P.		
PR	28-MAY-1999;	99US-0136782P.		
PR	01-JUN-1999;	99US-0137222P.		
PR	03-JUN-1999;	99US-0137528P.		
PR	04-JUN-1999;	99US-0137502P.		
PR	07-JUN-1999;	99US-0137724P.		
PR	08-JUN-1999;	99US-0138094P.		
PR	10-JUN-1999;	99US-0138540P.		
PR	10-JUN-1999;	99US-0138847P.		
PR	14-JUN-1999;	99US-0139119P.		

PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140891P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142877P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 94.4%; Score 861; DB 3; Length 1056;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCAGCTATTCTCAACAAGCAAGAACCTTAACCAAAAAGAAACGTTACCGAGACA 60

Db 128 ATCAGCTATTCTCAACAAGCAAGAACCTTAACCAAAAAGAAACGTTACCGAGACA 187

QY 61 TATGAGAGGAACATCACTTCGTGTTAGTTCACAAACGCTTATCATGAGCGCTGGATCTG 120  
DB |||||  
QY 188 TATGAGAGGAACATCACTTCGTGTTAGTTCACAAACGCTTATCATGAGCGCTGGATCTG 247  
DB |||||  
QY 121 GTACAAGCTCAAGCCCTCCCTTGAATCAGCCGGCCACCGCGTTACTGCTGCAACTCGC 180  
DB |||||  
QY 248 GTACAAGCTCAAGCCCTCCCTTGAATCAGCCGGCCACCGCGTTACTGCTGCAACTCGC 307  
DB |||||  
QY 181 CGCTCCCGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 240  
DB |||||  
QY 308 CGCTCCCGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 367  
DB |||||  
QY 241 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 300  
DB |||||  
QY 368 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 427  
DB |||||  
QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAAGATTAAGGT 360  
DB |||||  
QY 428 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAAGATTAAGGT 487  
DB |||||  
QY 361 TCTTGTTCTCAACGCTTCTTCCCGGACACAAACCGCGCTTCTCAGCTTCTGGA 420  
DB |||||  
QY 488 TCTTGTTCTCAACGCTTCTTCCCGGACACAAACCGCGCTTCTCAGCTTCTGGA 547  
DB |||||  
QY 421 CAGTATATGAGATCGCTTGGAGGTTTGGAGATTTGAGATTTTCATCTCATGAACAAG 480  
DB |||||  
QY 548 CAGTATATGAGATCGCTTGGAGGTTTGGAGATTTGAGATTTTCATCTCATGAACAAG 607  
DB |||||  
QY 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA 540  
DB |||||  
QY 608 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA 667  
DB |||||  
QY 541 AAATGTCCCATAGAGATTACGAGCTGGCAAAAATGTTGCATAGCAAGGTCATTTT 600  
DB |||||  
QY 668 AAATGTCCCATAGAGATTACGAGCTGGCAAAAATGTTGCATAGCAAGGTCATTTT 727  
DB |||||  
QY 601 CACAGGATCTATCAAGAAAGAAAGTTTAGCGGGAAGGATATGTTTCGGTCAACG 660  
DB |||||  
QY 728 CACAGAGGATCTATCAAGAAAGAAAGTTTAGCGGGAAGGATATGTTTCGGTCAACG 787  
DB |||||  
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGAT 720  
DB |||||  
QY 788 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGAT 847  
DB |||||  
QY 721 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATACATGCTGATGCT 780  
DB |||||  
QY 848 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATACATGCTGATGCT 907  
DB |||||  
QY 781 CTCCAAACCCAAAACCTTTTGACTCTCTCTGCTATTGCCCGATTATATGTAATA 840  
DB |||||  
QY 908 CTCCAAACCCAAAACCTTTTGACTCTCTCTGCTATTGCCCGATTATATGTAATA 967  
DB |||||  
QY 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900  
DB |||||  
QY 968 ATCTTAAGTCGGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 1027  
DB |||||  
QY 901 GCAACTTTCATC 912  
DB |||||  
QY 1028 GCAACTTTCATC 1039  
DB |||||

RESULT 3  
AAC47800  
ID AAC47800 standard; DNA; 909 BP.  
XX  
AC AAC47800;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55168.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-01231180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126284P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.



PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144086P.
PR	16-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	23-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	25-AUG-1999;	99US-0149930P.
PR	26-AUG-1999;	99US-0150566P.
PR	27-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157553P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160985P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 92.5%; Score 844; DB 3; Length 909;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 844; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	4	ACGCTATTCTCAAAACAAGCAAGAACCTTTAACCAAAAAGAAAACCTCTACGGAGACATAT 63
Db	10	ACGCTATTCTCAAAACAAGCAAGAACCTTTAACCAAAAAGAAAACCTCTACGGAGACATAT 69
Qy	64	GGAGAGGAAACATCACTTCGTGTTAGTTCAACAACGCTTATCATGGAGCCTGGATCTGGTA 123
Db	70	GGAGAGGAAACATCACTTCGTGTTAGTTCAACAACGCTTATCATGGAGCCTGGATCTGGTA 129
Qy	124	CAAGCTCAAGCCCTCCTTTGAATACGCGGCGCACCGGCTTACTGTGCGAATCTGCCCGC 183
Db	130	CAAGCTCAAGCCCTCCTTTGAATACGCGGCGCACCGGCTTACTGTGCGAATCTGCCCGC 189
Qy	184	CTCCGGGATCGACCCACGACCAATCCAGCCGTTGAACCCGTCGACGAATCTCCAAACC 243
Db	190	CTCCGGGATCGACCCACGACCAATCCAGCCGTTGAACCCGTCGACGAATCTCCAAACC 249
Qy	244	GTTGATCGAAACCCCTCAAAATCTTCCAGAGAACGAAGAGGTAATCTGGTTGGATTTCAG 303
Db	250	GTTGATCGAAACCCCTCAAAATCTTCCAGAGAACGAAGAGGTAATCTGGTTGGATTTCAG 309
Qy	304	CTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAAGTTCT 363
Db	310	CTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAAGTTCT 369

QY 364 TGTGTTCTCAACGGCTTCTTGCCCGACACAACCCACGTCGCTTCTCACGTTCTGGACAA 423  
DB |||||  
QY 370 TGTGTTCTCAACGGCTTCTTGCCCGACACAACCCACGTCGCTTCTCACGTTCTGGACAA 429  
DB |||||  
QY 424 GPATATGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACAAAGAAA 483  
DB |||||  
QY 430 GPATATGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACAAAGAAA 489  
DB |||||  
QY 484 TGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGTCCTTTACCAAAA 543  
DB |||||  
QY 490 TGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGTCCTTTACCAAAA 549  
DB |||||  
QY 544 TTGTCCTCATAGGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTTTCAC 603  
DB |||||  
QY 604 AGAGGATCTATCAAGAAGAAAAGTTTAGCGAGGAAGATATGTTGCGTGCACGAGT 663  
DB |||||  
QY 610 AGAGGATCTATCAAGAAGAAAAGTTTAGCGAGGAAGATATGTTGCGTGCACGAGT 669  
DB |||||  
QY 664 TTACGTAATGAGTACTGAAGACAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGA 723  
DB |||||  
QY 670 TTAAGTAATGAGTACTGAAGACAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGA 729  
DB |||||  
QY 724 TAAATTCACGTCCTGAAAGTCTACGAGATCGATGCGGAGATCACATGCTGATCTCTC 783  
DB |||||  
QY 730 TAAATTCACGTCCTGAAAGTCTACGAGATCGATGCGGAGATCACATGCTGATCTCTC 789  
DB |||||  
QY 784 CAAACCCCAAAACCTTTTGACTCTCTCTGCTATTGCGACCGATTATATGTAATAATC 843  
DB |||||  
QY 790 CAAACCCCAAAACCTTTTGACTCTCTCTGCTATTGCGACCGATTATATGTAATAATC 849  
DB |||||  
QY 844 TTA 847  
DB |||||  
QY 850 TTA 853  
DB |||||  
RESULT 4  
ADE82083  
ID ADE82083 standard; cdNA; 483 BP.  
XX AC ADE82083;  
XX DT 29-JAN-2004 (first entry)  
XX DE Arabidopsis thaliana expressed polynucleotide seq id 854.  
XX KW genetically modified organism; transgenic organism; plant;  
KW inhibitor testing; activator testing; modifier testing; fungicide;  
KW insecticide; genetic function; genetic regulation; cellular metabolism;  
KW gene; ss.  
XX OS Arabidopsis thaliana.  
XX PN US2003115639-A1.  
XX PD 19-JUN-2003.  
XX PF 26-JAN-2001; 2001US-00770961.  
XX PR 27-JAN-2000; 2000US-0178466P.  
XX (GORLACH J.  
PA (ANYI/) AN Y.  
PA (HAMI/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUYI/) YU Y.  
PA (RAME/) RAMEAKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOBS/) WOESSNER J P.

PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRIC/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX GORLACH J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX WPI; 2003-810930/76.  
XX Novel Arabidopsis thaliana nucleic acids useful for generating  
PT genetically modified transgenic organisms, for screening biologically  
PT active agents such as fungicides, insecticides.  
XX Claim 1; SEQ ID NO 854; 44pp; English.  
XX The invention describes a nucleic acid (I) comprising a sequence capable  
CC of hybridising under stringent conditions to any one of 999 fully defined  
CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
CC also useful for generating genetically modified and transgenic organisms,  
CC usually plant cells and plants. A protein encoded by (I) is useful in  
CC screening assays to determine the effect of candidate inhibitors,  
CC activators or modifiers of the gene product. The protein is also useful  
CC for screening biologically active agents e.g., fungicides and  
CC insecticides. A genetically modified cell, comprising an exogenous  
CC nucleic acid, where the nucleic acid comprises transcription regulatory  
CC sequences operably linked to a sequence capable of hybridising under  
CC stringent conditions to (I) is useful in the study of genetic function  
CC and regulation, for alteration of the cellular metabolism and for  
CC screening compounds that may affect the biological function of the gene  
CC or gene product. This sequence represents an Arabidopsis thaliana  
CC polynucleotide of the invention.  
XX SQ Sequence 483 BP; 129 A; 127 C; 112 G; 115 T; 0 U; 0 Other;  
Query Match 47.4%; Score 432; DB 10; Length 483;  
Best Local Similarity 99.8%; Pred. No. 4.3e-214;  
Matches 482; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 26 GAACCTTAACCAAAAGAAAACGCTACGGAGACATATGGAGAGAAACATCATTCTG 85  
DB 1 GAACCTTAACCAAAAGAAAACGCTACGGAGACATATGGAGAGAAACATCATTCTG 60  
QY 86 TTAGTTCAACACGCTTATCATGGAGCCCTGGATCTGGTCAAGCTCAAGCCCTCTTGA 145  
DB 61 TTAGTTCAACACGCTTATCATGGAGCCCTGGATCTGGTCAAGCTCAAGCCCTCTTGA 120  
QY 146 TCAGCCGGCCACCGGTTACTCTCGAACTCCGCCCTCCGGATCGACCCAGCACCA 205  
DB 121 TCAGCCGGCCACCGGTTACTCTCGAACTCCGCCCTCCGGATCGACCCAGCACCA 180  
QY 206 ATCCAGGCGGTTGAAACCGGTGCAAGATACCTCCAAACCGTTGATCGAAACCTTCA 265  
DB 181 ATCCAGGCGGTTGAAACCGGTGCAAGATACCTCCAAACCGTTGATCGAAACCTTCA 240  
QY 266 CTTCCAGAGAACGAGAGGTAATTCGTTGGATTCAGCTTCGGAGGATCAACATCGT 325  
DB 241 CTTTCAGAGAACGAGAGGTAATTCGTTGGATTCAGCTTCGGAGGATCAACATCGT 300  
QY 326 CTCGCCGGCGACATATTTCCGCGGAAGATTAAAGTTCTTGTTCTCCTCAAGCCCTTCT 385  
DB 301 CTCGCCGGCGACATATTTCCGCGGAAGATTAAAGTTCTTGTTCTCCTCAAGCCCTTCT 360  
QY 386 CCCGACACAAACCCAGCTGCTTCTCACGTTCTTGGCAAGATATATGGAGATGCTCGAGT 445  
DB |||||

Db 361 CCGACACACCCACGCTCTCTCACGTTCTGGACAAGTATATGAGATGCGCTGGAGGT 420  
 QY 446 TTGGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGAG 505  
 Db 421 TTGGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGAG 480  
 QY 506 ATG 508  
 Db 481 ATG 483

# RESULT 5 ABL33982/c

ID ABL33982 standard; DNA; 8592 BP.  
 XX  
 AC ABL33982;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1955.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antarteriosclerotic; anianaemic; cytosatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 KW ds.

OS Homo sapiens.

PN WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.

XX Claim 1; SEQ ID NO 1955; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention

XX Sequence 8592 BP; 2473 A; 132 C; 1657 G; 4330 T; 0 U; 0 Other;

Query Match 2.2%; Score 20; DB 6; Length 8592;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 CTCCTAACCCCAAACTCT 800

Db 2474 CTCCTAACCCCAAACTCT 2455

# RESULT 6

AAI80037/c

ID AAI80037 standard; cDNA; 371 BP.

XX AAI80037;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 97.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR P-PSDB; AAO00106.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX Claim 1; SEQ ID NO 97; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 371 BP; 104 A; 43 C; 55 G; 169 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 4; Length 371;

Best Local Similarity 100.0%; Pred. No. 48;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 CATCTCATGAAACAAGAAA 483

Db 275 CATCTCATGAAACAAGAAA 257

# RESULT 7

ADD49350/c

ID ADD49350 standard; cDNA; 387 BP.

XX ADD49350;

XX 15-JAN-2004 (first entry)

XX Human lung cancer associated cDNA 61483100.



KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
 XX stress resistance.  
 XX Gossypium hirsutum.  
 XX US2004181830-A1.  
 XX 16-SEP-2004.  
 XX 29-JAN-2004; 2004US-00767795.  
 XX 07-MAY-2001; 2001US-00849529.  
 PR 12-DEC-2001; 2001US-00021323.  
 XX (KOVA/) KOVALIC D K.  
 PA (ZHOU/) ZHOU Y.  
 PA (CAO/) CAO Y.  
 XX Kovalic DK, Zhou Y, Cao Y;  
 PI WPI; 2004-667718/65.  
 XX New recombinant nucleic acid molecules and polypeptides from Gossypium  
 PT hirsutum, useful for producing plants with improved biological  
 PT characteristics (e.g. improved plant cold or drought tolerance).  
 XX Claim 1; SEQ ID NO 5332; 14pp; English.  
 XX The invention relates to a recombinant polynucleotide comprising any of  
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.  
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
 CC sequences mentioned in the specification and producing a plant having an  
 CC improved property. Producing a plant having an improved property  
 CC comprises transforming a plant with a recombinant construct comprising a  
 CC promoter region functional in a plant cell operably joined to a  
 CC polynucleotide comprising a coding sequence for a polypeptide associated  
 CC with the property, and growing the transformed plant. The polypeptide is  
 CC useful for improving plant cold tolerance, manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, improving plant  
 CC drought tolerance, providing increased resistance to plant disease,  
 CC producing galactanmannan (or lignin or plant growth regulators), improving  
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
 CC the rate of homologous recombination in plants, improving plant tolerance  
 CC to extreme osmotic conditions or to pathogens or pests, improving yield  
 CC by modification of photosynthesis, modifying seed oil or protein yield  
 CC and/or content, improving yield by modification of carbohydrate, nitrogen  
 CC or phosphorus use and/or uptake, or improving yield by providing improved  
 CC plant growth and development under at least one stress condition. The  
 CC polynucleotide and polypeptide may also be used in recombinant DNA  
 CC constructs, in physical arrays of molecules, as plant breeding markers,  
 CC or in computer-based storage and analysis systems. The present sequence  
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequences.html?DocID=20040181830. However only 6585  
 CC polynucleotide sequences were available, the remaining 5213  
 CC polynucleotides and all 58798 protein sequences were not present.  
 XX Sequence 432 BP; 127 A; 81 C; 79 G; 145 T; 0 U; 0 Other;  
 SQ

Query Match 2.1%; Score 19; DB 13; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 880 AAACAAACCCCTTTTTC 898  
 DB 1 AAACAAACCCCTTTTTC 19

RESULT 10  
 ADD49911/c  
 ID ADD49911 standard; cDNA; 440 BP.  
 XX

AC ADD49911;  
 XX 15-JAN-2004 (first entry)  
 XX Human lung cancer associated cDNA 61824027.  
 XX Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;  
 KW vaccine; T-cell; tumour.  
 XX Homo sapiens.  
 XX US2003194764-A1.  
 XX 16-OCT-2003.  
 XX 04-APR-2002; 2002US-00116712.  
 XX 05-APR-2001; 2001US-0282289P.  
 PR 05-OCT-2001; 2001US-0327511P.  
 XX (CORI-) CORIXA CORP.  
 XX Bangur CS, Switzer A;  
 XX WPI; 2003-844452/78.  
 XX New isolated polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cancer, particularly lung cancer.  
 PT Claim 1; SEQ ID NO 643; 250pp; English.  
 XX The invention relates to an isolated polynucleotide (a) comprising any of  
 CC the 666 fully defined nucleotide sequences appearing as ADD49269 -  
 CC ADD49935, ADD49936 and ADD49938, complements of (a); sequences of at  
 CC least 20 contiguous residues of (a); sequences that hybridise to (a)  
 CC under highly stringent conditions; sequences having at least 75 or 90%  
 CC identity to (a); or degenerate variants of (a). Also included are an  
 CC isolated polypeptide (b) (comprising: sequences encoded by the new  
 CC polynucleotide; any of the 4 amino acid sequences fully defined in the  
 CC specification; or sequences having at least 70 or 90% identity to the  
 CC sequence in (a) or (b)), an expression vector comprising the above  
 CC polynucleotide operably linked to an expression control sequence, a host  
 CC cell transformed or transfected with the above expression vector, an  
 CC isolated antibody, or its antigen-binding fragment, that specifically  
 CC binds to the above polypeptide, an oligonucleotide that hybridises to the  
 CC above-mentioned nucleotide sequences under highly stringent conditions, a  
 CC fusion protein comprising at least one polypeptide cited above, detecting  
 CC the presence of a cancer in a patient (comprising: obtaining a biological  
 CC sample from the patient; contacting the biological sample with a binding  
 CC agent that binds to the polypeptide, or with the oligonucleotide cited  
 CC above; detecting in the sample an amount of the polypeptide that binds to  
 CC the binding agent, or an amount of a polynucleotide that hybridises to  
 CC the oligonucleotide; and comparing the amount of polypeptide, or  
 CC polynucleotide that hybridises to the oligonucleotide, to a predetermined  
 CC cut-off value and then determining the presence of a cancer in the  
 CC patient), a method for stimulating and/or expanding T-cells specific for  
 CC a tumour protein (comprising contacting T-cells with the above  
 CC polypeptide, polynucleotide or antigen-presenting cells that express the  
 CC polypeptide, under conditions and for a time sufficient to permit the  
 CC stimulation and/or expansion of T-cells), an isolated T-cell population  
 CC comprising T-cells prepared by the method, a composition comprising a  
 CC first component selected from physiological carriers and  
 CC immunostimulants, and a second component selected from the above  
 CC polypeptide, polynucleotide, antibody, fusion protein, T-cell population  
 CC and antigen-presenting cells that express the above polypeptide,  
 CC stimulating an immune response in a patient (comprising administering to  
 CC the patient the above composition) treating lung cancer in a patient  
 CC (comprising administering to the patient the above composition and a  
 CC diagnostic kit (comprising: at least one oligonucleotide cited above; or  
 CC at least one antibody cited above and a detection reagent, where the  
 CC detection reagent comprises a reporter group). The composition and  
 CC methods are useful in diagnosing, preventing and treating cancer,  
 CC particularly lung cancer. The present sequence is a lung cancer-

```
CC associated antigen cDNA of the invention.
XX
SQ Sequence 440 BP; 144 A; 61 C; 122 G; 113 T; 0 U; 0 Other;
    Query Match      2.1%; Score 19; DB 10; Length 440;
    Best Local Similarity 100.0%; Pred. No. 48;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGCAACTTT 908
Db 137 CCTTTTCCGGCAACTTT 119

RESULT 11
ABK39793/c
ID ABK39793 standard; cDNA; 624 BP.
XX
AC ABK39793;
XX
XX 21-MAY-2002 (first entry)
XX
XX DNA encoding recombinant protein L548S.
XX
XX Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
XX Gene; ss.
XX
XX Homo sapiens.
XX
XX WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US022058.
XX
XX 11-JUL-2000; 2000US-00614124.
XX
XX 29-AUG-2000; 2000US-00651563.
XX
XX 08-SEP-2000; 2000US-00658824.
XX
XX 26-SEP-2000; 2000US-00671325.
XX
XX 06-OCT-2000; 2000US-00677419.
XX
XX 30-OCT-2000; 2000US-00702705.
XX
XX 13-DEC-2000; 2000US-00736457.
XX
XX 03-MAY-2001; 2001US-00849626.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
PI Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI; 2002-164634/31.
XX
XX P-PSDB; AAU85606.
XX
XX Novel polynucleotide encoding a lung tumor polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumor protein.
XX
XX Claim 9; SEQ ID NO 1914; 223pp; English.
XX
XX The invention describes an isolated polynucleotide and polypeptide useful
CC for stimulating and/or expanding T cells specific for a tumour protein
CC for determining the presence of a cancer in a patient. A composition
CC containing the polynucleotide and/or polypeptide is useful for treating a
CC lung cancer in a patient. The polypeptide is useful for removing tumour
CC cells from a biological sample. The polynucleotide is also useful as
CC probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence encodes a lung tumour associated protein or protein
CC fragment, described in the method of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;
SQ
    Query Match      2.1%; Score 19; DB 6; Length 624;

Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGCAACTTT 908
Db 559 CCTTTTCCGGCAACTTT 541

RESULT 12
ACA12122/c
ID ACA12122 standard; cDNA; 624 BP.
XX
XX ACA12122;
XX
XX 06-JUN-2003 (first entry)
XX
XX Human His tagged lung cancer associated protein L548S, cDNA.
XX
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
XX T cell expansion; CD4; CD8; RAI2; Gene.
XX
XX Homo sapiens.
XX
XX US2002197669-A1.
XX
XX 26-DEC-2002.
XX
XX 03-MAY-2001; 2001US-00849626.
XX
XX 13-DEC-2000; 2000US-00736457.
XX
XX (BANG/) BANGUR C S.
XX
XX (FANG/) FANGER G R.
XX
XX (WANG/) WANG T.
XX
XX (WANG/) WANG T.
XX
XX (SWIT/) SWITZER A P.
XX
XX (MCNE/) MCNEILL P D.
XX
XX (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX
XX WPI; 2003-352750/33.
XX
XX P-PSDB; ABU69588.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
XX compositions, e.g. vaccines, for treating lung cancer.
XX
XX Example 14; Page; 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
```

CC presenting cells that express the polynucleotide, such that T cells  
CC proliferate, administering to the patient an effective amount of the  
CC proliferated T cells, and thus inhibiting the development of a cancer in  
CC the patient. The polynucleotide, protein and cells are useful in a  
CC composition for stimulating an immune response in a patient, and for  
CC treating a cancer in a patient (particularly lung cancer). The  
CC oligonucleotide is useful for determining the presence of a cancer in a  
CC patient. The protein and oligonucleotides are useful in pharmaceutical  
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
CC or primer for nucleic acid hybridisation, and in the design and  
CC preparation of ribozyme molecules for inhibiting expression of tumour  
CC polypeptides and proteins in tumour cells. An amplified portion of the  
CC polynucleotide is useful for isolating a full-length gene from a suitable  
CC library. The present sequence encodes a fusion protein of human RA12 with  
CC the protein product of a cDNA (full length, extended or partial) isolated  
CC from a library derived from lung tumour/cancer cells. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from the USPTO at  
CC seqdata.uspto.gov/sequence.html?docId=20020197669  
XX  
SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 8; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTTCGGGCAACTTT 908  
DB 559 CCTTTTTCGGGCAACTTT 541

RESULT 13  
ACA03308/c  
ID ACA03308 standard; DNA; 624 BP.  
XX  
AC ACA03308;  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE Lung cancer therapy and diagnosis associated DNA #15.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ds.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-00902941.

XX 30-JUN-1999; 99US-00346492.

XX 15-OCT-1999; 99US-00419356.

XX 17-DEC-1999; 99US-00466867.

XX 30-DEC-1999; 99US-00476300.

XX 06-MAR-2000; 2000US-00519642.

XX 22-MAR-2000; 2000US-00533077.

XX 10-APR-2000; 2000US-00546259.

XX 27-APR-2000; 2000US-00560406.

XX 05-JUN-2000; 2000US-00589184.

XX 11-JUL-2000; 2000US-00614124.

XX 29-AUG-2000; 2000US-00651563.

XX 08-SEP-2000; 2000US-00658824.

XX 26-SEP-2000; 2000US-00671325.

XX 06-OCT-2000; 2000US-00677419.

XX 30-OCT-2000; 2000US-00702705.  
XX 13-DEC-2000; 2000US-00736457.  
XX 03-MAY-2001; 2001US-00849626.  
XX  
XX (CORI-) CORIXA CORP.  
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;  
XX Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.  
XX  
XX New polynucleotide, useful for preparing a composition for treating or  
XX inhibiting development of cancer, e.g. lung cancer.  
XX

XX Example 14; SEQ ID NO 1914; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
XX sequences, complement or degenerate variants of them. The polynucleotide  
XX is useful for preparing a composition e.g. a vaccine or for gene therapy,  
XX for treating or inhibiting development of cancer, e.g. lung cancer. This  
XX sequence represents a polynucleotide associated with the compositions and  
XX methods for the therapy and diagnosis of lung cancer

XX Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 8; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTTCGGGCAACTTT 908

DB 559 CCTTTTTCGGGCAACTTT 541

RESULT 14  
ADH47433/c  
ID ADH47433 standard; DNA; 624 BP.  
XX  
AC ADH47433;  
XX  
DT 25-MAR-2004 (first entry)  
XX  
DE Human lung tumour DNA, SEQ ID No 1914.

XX lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;  
XX human; ds.

XX Homo sapiens.

XX WO2003037267-A2.

XX 08-MAY-2003.

XX 28-OCT-2002; 2002WO-US034777.

XX 29-OCT-2001; 2001US-00017754.

XX 28-MAR-2002; 2002US-00113872.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;  
XX Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;  
XX Bangur CS, McNabb A;

XX WPI; 2003-468346/44.

XX New polypeptides and encoding polynucleotides, useful for diagnosing,  
XX preventing and/or treating lung cancer.

XX Example 14; SEQ ID NO 1914; 258pp; English.

XX The invention relates to novel compositions and methods for the therapy  
XX and diagnosis of cancer, particularly lung cancer. The compositions  
XX comprise one or more lung tumour polypeptides, immunogenic portions  
XX thereof, polynucleotides that encode such polypeptides, antigen presenting  
XX cells that express such polypeptides, and T cells that are specific for  
XX cells expressing such polypeptides. The novel compositions have  
XX cytostatic and immunostimulant activity. The lung tumour antigens can be  
XX used in the creation of a vaccine. The polynucleotides that encode the  
XX lung tumour polypeptides can be used in gene therapy to help in the  
XX treatment of lung tumours. This polynucleotide represents a human lung

CC tumour DNA sequence of the invention. This sequence was not shown in the  
CC specification. It has been taken from a World Intellectual Property  
CC Organization CD ROM supplied with the specification.

XX  
SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;  
Query Match 2.1%; Score 19; DB 10; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 890 CCTTTTTCGGGCAACTTT 908  
|||||  
Db 559 CCTTTTTCGGGCAACTTT 541

RESULT 15  
ADJ21352/c  
ID ADJ21352 standard; DNA; 624 BP.  
XX  
AC ADJ21352;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human lung cancer-related L548S coding region DNA.  
XX  
KW lung tumour; cytostatic; lung cancer; human; ds; gene; L548S.

XX  
OS Homo sapiens.  
XX  
PN US2003211510-A1.  
XX  
PD 13-NOV-2003.

XX  
PF 28-OCT-2002; 2002US-00283017.  
XX  
PR 30-JUN-1999; 99US-00346492.  
PR 15-OCT-1999; 99US-00419356.  
PR 17-DEC-1999; 99US-00466867.  
PR 30-DEC-1999; 99US-00476300.  
PR 06-MAR-2000; 2000US-00519642.  
PR 22-MAR-2000; 2000US-00533077.  
PR 10-APR-2000; 2000US-00546259.  
PR 27-APR-2000; 2000US-00560406.  
PR 05-JUN-2000; 2000US-00589184.  
PR 11-JUL-2000; 2000US-00614124.  
PR 29-AUG-2000; 2000US-00651563.  
PR 08-SEP-2000; 2000US-00658824.  
PR 26-SEP-2000; 2000US-00671325.  
PR 06-OCT-2000; 2000US-00677419.  
PR 30-OCT-2000; 2000US-00702705.  
PR 13-DEC-2000; 2000US-00736457.  
PR 03-MAY-2001; 2001US-00849626.  
PR 10-JUL-2001; 2001US-00902941.  
PR 29-OCT-2001; 2001US-00017754.  
PR 28-MAR-2002; 2002US-00113872.

XX (CORI-) CORIXA CORP.  
XX  
PI Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;  
PI Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;  
PI Bangur CS, McNabb A;  
XX  
DR WPI; 2004-167010/16.  
DR P-PSDB; ADJ21351.

XX  
PT Novel polynucleotide encoding lung tumor polypeptides, useful for  
PT diagnosing, preventing and treating cancer e.g. lung cancer.

XX  
PS Example 14; SEQ ID NO 1914; 99pp; English.

XX  
CC The invention relates to a novel isolated polynucleotide comprising a  
CC sequence chosen from any one of 40 lung tumour polypeptides or its  
CC complements, fragments or degenerate variants. The method of the

CC invention has cytostatic applications and may be useful for detecting and  
CC treating lung cancer in a patient, as well as for inhibiting the  
CC development of lung cancer in a patient via incubating CD4+ and/or CD8+ T  
CC cells isolated from a patient with at least one component chosen from a  
CC polypeptide, polynucleotide or antigen presenting cell (APC) of the  
CC invention and administering an effective amount of the proliferated T  
CC cells to the patient. The current sequence is that of the human lung  
CC cancer-related DNA of the invention. The current sequence is not shown in  
CC the specification per se but is available on the USPTO web-site  
CC http://seqdata.uspto.gov/sequence.html?DocID=20030211510.

SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;  
Query Match 2.1%; Score 19; DB 13; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 890 CCTTTTTCGGGCAACTTT 908  
|||||  
Db 559 CCTTTTTCGGGCAACTTT 541

Search completed: September 12, 2005, 13:13:17  
Job time : 505.287 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 06:38:26 ; Search time 3957.19 Seconds  
(without alignments)  
11167.303 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcacgctattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_scs:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	94.4	928	8 AY046015	AY046015 Arabidops
2	808	88.6	808	8 AY142490	AY142490 Arabidops
3	777	85.2	777	8 AY093714	AY093714 Arabidops
4	638	70.0	911	8 AY058115	AY058115 Arabidops
5	424	46.5	57180	8 APT18022	AL360334 Arabidops
6	140	15.4	856	12 AY202287	AY202287 Arabidops
7	133	14.6	208	11 BX321957	BX321957 Arabidops
8	35	3.8	160	11 AL773330	AL773330 Arabidops
9	23	2.5	961	8 AF178576	AF178576 Ruvolfia
10	23	2.5	238185	2 AC111417	AC111417 Rattus no
11	22	2.4	143168	8 AP003765	AP003765 Oryza sat
12	22	2.4	146690	8 AP004671	AP004671 Oryza sat
13	22	2.4	163246	9 AC010196	AC010196 Homo sapi
14	22	2.4	187111	9 AC087244	AC087244 Homo sapi
15	21	2.3	31740	2 AC025128	AC025128 Homo sapi
16	21	2.3	39210	3 U41508	U41508 Caenorhabdi
17	21	2.3	101620	9 AC102942	AC102942 Homo sapi
18	21	2.3	139301	8 AP003537	AP003537 Oryza sat
19	21	2.3	143739	8 AC132483	AC132483 Oryza sat

20	21	2.3 147588	9 AC098612	AC098612 Homo sapi
21	21	2.3 161850	9 AF306443	AF306443 Homo sapi
c 22	21	2.3 235029	2 AC131419	AC131419 Rattus no
c 23	20	2.2 3819	5 BC070706	BC070706 Xenopus l
c 24	20	2.2 8592	6 AX346884	AX346884 Sequence
25	20	2.2 12042	1 AB001089	AB001089 Archaeogl
26	20	2.2 43351	3 CBRG47121	AC084674 Caenorhab
c 27	20	2.2 54252	8 AB017060	AB017060 Arabidops
c 28	20	2.2 66805	2 AC101558	AC101558 Mus muscu
c 29	20	2.2 110000	2 AP006491_3	Continuation (4 of
c 30	20	2.2 110121	9 AL590989	AL590989 Human DNA
31	20	2.2 134701	5 BX255885	BX255885 Zebrafish
32	20	2.2 167619	5 AL929045	AL929045 Zebrafish
c 33	20	2.2 174569	9 AC022533	AC022533 Homo sapi
c 34	20	2.2 175171	10 AC112158	AC112158 Mus muscu
c 35	20	2.2 176155	2 AC020879	AC020879 Mus muscu
c 36	20	2.2 176775	2 AC016978	AC016978 Homo sapi
37	20	2.2 178968	9 AF006274	AF006274 Papio ham
38	20	2.2 192808	10 AL645566	AL645566 Mouse DNA
39	20	2.2 195918	10 AL805935	AL805935 Mouse DNA
c 40	20	2.2 199885	2 AC116869	AC116869 Mus muscu
41	20	2.2 210472	9 AC035139	AC035139 Homo sapi
42	20	2.2 218807	2 BX511249	BX511249 Mus muscu
43	20	2.2 231764	5 BX005426	BX005426 Zebrafish
44	20	2.2 240691	2 AC108297	AC108297 Rattus no
45	20	2.2 241421	2 AC110427	AC110427 Rattus no

#### ALIGNMENTS

AY046015 928 bp mRNA linear PLN 18-SEP-2002  
Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
(AF5910300) mRNA, complete cds.

ACCESSION AY046015.1 GI:15028130

VERSION FLI CDNA.

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 928)

AUTHORS Yanada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 928)

Yanada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Miranda,M., Narusaka,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Direct Submission

Submitted (06-JUL-2001) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,

Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Barth, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

Location/Qualifiers

1..928

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAFL05-09-N03 (R09923)"

/ecotype="Columbia"

/note="this clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."

1..928

/gene="At5g10300"

1..61

/gene="At5g10300"

62..838

/gene="At5g10300"

/codon\_start=1

/evidence="experimental"

/product="putative alpha-hydroxynitrile lyase"

/protein\_id="AAK76689.1"

/db\_xref="GI:15028131"

/translation="MERKHFVIVHNAVGVWYKLPLESAGHRVTAVELAASGI  
DRPQAVETVDEYSKPLIETLKSIPENEVILVSGFSGGINIALAADI PAKIKVLV  
FLNAPLPTTHVPSHLVDKYMPPGLGDCFSFHSHTRNGLMKGPKFMKRLQ  
NCPIEDYELAKMLHRQSPFTEDLSKEKFESEGYGSVQRYVVMSSDKAIPCFIRW  
MIDNFNVSKVYBIDGGDHVWMLSKPQKLPDLSALATDYM"

misc\_difference 590

/gene="At5g10300"

/note="compared to genomic sequence resulting in an amino acid sequence difference"

/replace="g"

839..928

/gene="At5g10300"

#### 3' UTR

#### ORIGIN

Query Match 94.4%; Score 861; DB 8; Length 928;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCAGCGTATTCTCAACAGCAAGAACCTTAACCAAAAGAAACCTCTACGGAGACA 60  
DB 1 ATCAGCGTATTCTCAACAGCAAGAACCTTAACCAAAAGAAACCTCTACGGAGACA 60

QY 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTACCAACGCTTATCATGAGCGCTGGATCTG 120  
DB 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTACCAACGCTTATCATGAGCGCTGGATCTG 120

QY 121 GTACAGAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCGTTACTGCTGCAACTCGC 180  
DB 121 GTACAGAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCGTTACTGCTGCAACTCGC 180

QY 181 CGCCTCCGGATCGACCCAGACCAATCCAGCGCTTGAACCGTCGACGAATCTCCAA 240  
DB 181 CGCCTCCGGATCGACCCAGACCAATCCAGCGCTTGAACCGTCGACGAATCTCCAA 240

QY 241 ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAACGAAGAGGTAATTCGTGTTGGATT 300  
DB 241 ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAACGAAGAGGTAATTCGTGTTGGATT 300

QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAGATTAAGGT 360  
DB 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAGATTAAGGT 360

QY 361 TCTTGTGTCTCTCAACGCCCTTCTTGGCCCGACACAAACCCACGTCGCTTCTCACGTCCTGGA 420  
DB 361 TCTTGTGTCTCTCAACGCCCTTCTTGGCCCGACACAAACCCACGTCGCTTCTCACGTCCTGGA 420

QY 421 CAAGTATATGAGATGCTGGAGGTTTGGGAGATGTTGAGTGTTCATCTCATGAACAAAG 480  
DB 421 CAAGTATATGAGATGCTGGAGGTTTGGGAGATGTTGAGTGTTCATCTCATGAACAAAG 480

QY 481 AAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGCACCTCTTTACCA 540  
DB 481 AAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGCACCTCTTTACCA 540

QY 541 AAATTTGTCCTCATGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT 600  
DB 541 AAATTTGTCCTCATGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT 600

QY 601 CACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGGTTCGGTGCACAG 660  
DB 601 CACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGGTTCGGTGCACAG 660

QY 661 AGTTTACGTAATGAGTAGTAGAAGCAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 720  
DB 661 AGTTTACGTAATGAGTAGTAGAAGCAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 720

QY 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780  
DB 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780

QY 781 CTCCAAAACCCCAAAAACCTCTTTGATCTCTCTCTGCTATTTGCCACCGATATATGTAATA 840  
DB 781 CTCCAAAACCCCAAAAACCTCTTTGATCTCTCTCTGCTATTTGCCACCGATATATGTAATA 840

QY 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTTACTATAAAACAAACCCCTTTTCCGG 900  
DB 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTTACTATAAAACAAACCCCTTTTCCGG 900

QY 901 GCAACTTTCATC 912  
DB 901 GCAACTTTCATC 912

RESULT 2  
LOCUS AY142490 808 bp mRNA linear PLN 23-SEP-2002  
DEFINITION Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
(At5g10300) mRNA, complete cds.  
ACCESSION AY142490  
VERSION AY142490.1 GI:23296321  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 808)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Arabidopsis Open Reading Frame (ORF) Clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 808)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Direct Submission

JOURNAL Submitted (16-AUG-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1. .808

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="C104719"

/ecotype="Columbia"

/notes="This clone is in pUNI 51."

gene

1. .808

/genes="At5g10300"

CDS

1. 777

/gene="At5g10300"

/codon\_start=1

/evidence=experimental

/product="putative alpha-hydroxynitrile lyase"

/protein\_id="AA13041.1"

/db\_xref="GI:23296322"

/translations="MERKHFVLVHNAYHGAWIYKLLPESAGHRVTVAVELASGI  
DRPIQAVTEVSEYKPLIETLKSLENBEVILVFGSGINIALAADIPPAKIKVLV  
FLNAPLDTHVPSHVDIKYEMPGGLGCEFSHETRNGLSLKMGPKFKARLYQ  
NCPIDYELAKMLHROGSPFTEDLSKKEFKSEGVGVVMSSEDAIPCFDIRW  
MIDNFNVSKVEIDGGDHVWMLSKPKQLFDSLAIATDYM"

3' UTR

778. .808

/genes="At5g10300"

ORIGIN

Query Match 88.6%; Score 808; DB 8; Length 808;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 808; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 ATGGAGGAGAAACATCATTCTGTTAGTTTCAACAGCTTATCATGAGCGCTGGATCTGG 121

DB 1 ATGGAGGAGAAACATCATTCTGTTAGTTTCAACAGCTTATCATGAGCGCTGGATCTGG 60

QY 122 TACAAGCTCAAGCCCTCTTGAATCAGCGCGCCAGCGGTACTGCTCGAATCTGGC 181

DB 61 TACAAGCTCAAGCCCTCTTGAATCAGCGCGCCAGCGGTACTGCTCGAATCTGGC 120

QY 182 GCCTCCGGGATCGACCGACCAATCCAGCGCGTTGAAACCGTTCGACGAATCTCCAAA 241

DB 121 GCCTCCGGGATCGACCGACCAATCCAGCGCGTTGAAACCGTTCGACGAATCTCCAAA 180

QY 242 CGTTGATCGAAACCTCAAATCTTTCAGAGAACGAGAGGTAAATCTGGTTGGATTC 301

DB 181 CGTTGATCGAAACCTCAAATCTTTCAGAGAACGAGAGGTAAATCTGGTTGGATTC 240

QY 302 AGCTTCGGAGGATCAACATCGTCTCGCGCGGCACATATTTCCGGGAGGATTAAGTT 361

DB 241 AGCTTCGGAGGATCAACATCGTCTCGCGCGGCACATATTTCCGGGAGGATTAAGTT 300

QY 362 CTGTGTCTCTCAACGCTTCTTCCCGCACACACCGTCTCTCAAGTTCTCGAC 421

DB 301 CTGTGTCTCTCAACGCTTCTTCCCGCACACACCGTCTCTCAAGTTCTCGAC 360

QY 422 AAGTATATGAGATGCTCGAGGTTTGGAGATTTGAGTTTTCATCTCATGAACAAGA 481

DB 361 AAGTATATGAGATGCTCGAGGTTTGGAGATTTGAGTTTTCATCTCATGAACAAGA 420

QY 482 AATGGACCATGAGTTTATGAGATGGGACCAAAATTCATGAAGCAGTCTTTACCAA 541

DB 421 AATGGACCATGAGTTTATGAGATGGGACCAAAATTCATGAAGCAGTCTTTACCAA 480

QY 542 AATTGTCCCATAGAGATTACGAGCTGGGCAAAATGTTCATAGCAAGGGTCATTTTC 601

DB 481 AATTGTCCCATAGAGATTACGAGCTGGGCAAAATGTTCATAGCAAGGGTCATTTTC 540

QY 602 ACAGAGGATCTATCAAGAGAAAGAAAGTTTACGAGGAGGATATGTTTCGTTGCAACGA 661

Db 541 ACAGAGGATCTATCAAGAGAAAGAAAGTTTACGAGGAGGATATGTTTCGTTGCAACGA 600

QY 662 GTTTACGTAATAGTAGTAGACACAAAGCCATCCCTCGATTTTCATTCGTTGGATGATT 721

Db 601 GTTTACGTAATAGTAGTAGACACAAAGCCATCCCTCGATTTTCATTCGTTGGATGATT 660

QY 722 GATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTGTGCTC 781

Db 661 GATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTGTGCTC 720

QY 782 TCCAAACCCCAAAAACCTCTTTCGACTCTCTCTGCTATTCGCCACCGATTATATGTAATA 841

Db 721 TCCAAACCCCAAAAACCTCTTTCGACTCTCTCTGCTATTCGCCACCGATTATATGTAATA 780

QY 842 TCTTAAGTCGCTTTTACTTTTTTTCAT 869

Db 781 TCTTAAGTCGCTTTTACTTTTTTTCAT 808

RESULT 3

AY093714

LOCUS AY093714

DEFINITION Arabidopsis thaliana AT5g10300/F19D22\_70 mRNA, complete cds.

ACCESSION AY093714

VERSION AY093714.1

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 777)

AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submission

JOURNAL Submitted (29-MAR-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)



/product="AT5g10300/F18D22\_70"  
/protein\_id="AAL25532.1"  
/db\_xref="GI:16648679"  
/translation="MERKHFVLVNHAYHGAMIWYKLPILLSAGHRVTAVLAASGI  
DRPQAVDEYSHVSKPLLETUKSQENBEGILVFPFGINIALAADLFPAPKIKLVY  
FLNAPUDTTHVPSHVLVDKYMMPGGLGDCFFSSHBTRNGTWSLLKMGPKFKARLYQ  
NCPIDYELAKMLHRQGSFPTEDLSKKEKFSBEGSVQRVVTVMSSDKAIPCDFIRW  
MIDNFVSKVYEIDGGDHVMLS KPQLFDSLAIATDM"  
847. .911

ORIGIN	3'UTR	847. .911
Query Match		
Best Local Similarity 99.5%; Score 638; DB 8; Length 911;		
Matches 838; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
QY	61	TATGGAGAGAAACATCACTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 120
DB	69	TATGGAGAGAAACATCACTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG 128
QY	121	GTACAAGCTCAAGCCCTCTTGAATCAGCCGCCACCGCTTACTGCTGTGAACTCGC 180
DB	129	GTACAAGCTCAAGCCCTCTTGAATCAGCCGCCACCGCTTACTGCTGTGAACTCGC 188
QY	181	CGCTCCGGATCGACCCACGACCAATCCAGGCCGTTGAAACCGTTCGACGAATATCTCAA 240
DB	189	CGCTCCGGATCGACCCACGACCAATCCAGGCCGTTGAAACCGTTCGACGAATATCTCAA 248
QY	241	ACGTTGATCGAAACCTCAATCTCTTCAGAGAACGAGGATTAATCTCGTGGATT 300
DB	249	ACGTTGATCGAAACCTCAATCTCTTCAGAGAACGAGGATTAATCTCGTGGATT 308
QY	301	CAGTTCCGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGT 360
DB	309	CAGTTCCGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGT 368
QY	361	TCTTGTGTTCTCAACGCCTTCTTGCCCGACACCAACCCACGTCCTTCACGTTCTGGA 420
DB	369	TCTTGTGTTCTCAACGCCTTCTTGCCCGACACCAACCCACGTCCTTCACGTTCTGGA 428
QY	421	CAAGTATATGAGATCGCTGGAGTTTGGAGATTGTGAGTTTCAATCATGAAACAG 480
DB	429	CAAGTATATGAGATCGCTGGAGTTTGGAGATTGTGAGTTTCAATCATGAAACAG 488
QY	481	AAATGGACGATGAGTTTATGAAGATGGACCAAAATTCATGAAGCACGCTTTACCA 540
DB	489	AAATGGACGATGAGTTTATGAAGATGGACCAAAATTCATGAAGCACGCTTTACCA 548
QY	541	AAATTGCTCCATAGAGATTACGAGCTGCGCAAAATGTTTGCATAGCAAGGTCATTTT 600
DB	549	AAATTGCTCCATAGAGATTACGAGCTGCGCAAAATGTTTGCATAGCAAGGTCATTTT 608
QY	601	CACAGAGATCTATCAAGAGAAAGTTTACGAGGAAGATATGTTTCGGTCAACG 660
DB	609	CACAGAGATCTATCAAGAGAAAGTTTACGAGGAAGATATGTTTCGGTCAACG 668
QY	661	AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGATGAT 720
DB	669	AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGATGAT 728
QY	721	TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTATGCT 780
DB	729	TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTATGCT 788
QY	781	CTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATGTCACCGGATATATGTAATA 840
DB	789	CTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATGTCACCGGATATATGTAATA 848
QY	841	ATCTTAAGTCGTTTACTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGG 900
DB	849	ATCTTAAGTCGTTTACTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGG 908
QY	901	GC 902

Db	909 GC 910
RESULT 5	
ATF18D22	
LOCUS	57180 bp DNA linear PLN 05-JUL-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 5, BAC clone F18D22 (BSSA project).
ACCESSION	AL360334
VERSION	AL360334.1
KEYWORDS	GI:8953404
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	Bevan, M., Zimmermann, W., Grueniseisen, A., Wambutt, R., Kalicki, J., Wohldmann, P., Smith, A., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.
JOURNAL	Unpublished
REFERENCE	2 (bases 52233 to 57180)
AUTHORS	Bevan, M., Voickaert, G., Grymonprez, B., Voet, M., Robben, J., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 57180)
AUTHORS	EU Arabidopsis sequencing, project.
TITLE	Direct Submission
JOURNAL	Submitted (05-JUL-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES	Location/Qualifiers
source	1..57180
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="5"
misc_feature	1..20186
	/note="overlap to BAC T31P16, please refer to
gene	EMBL:AL356332 for analysis and annotation"
	17729..21213
gene	/gene="F18D22_10"
	complement(join(17729..17941,18038..18145,18237..18323,18409..18489,18699..18833,18917..19138,19223..19303,19389..19550,19645..19739,19924..20019,20122..20263,20453..20548,20735..20873,21134..21213))
CDS	/gene="F18D22_10"
	complement(join(17729..17941,18038..18145,18237..18323,18409..18489,18699..18833,18917..19138,19223..19303,19389..19550,19645..19739,19924..20019,20122..20263,20453..20548,20735..20873,21134..21213))
	/gene="F18D22_10"
	/note="Contains Glutamine amidotransferases class-II active site AA1-6"
	/codon_start=1
	/product="asparagine synthetase ASN3"
	/protein_id="CAB96680.1"
	/db_xref="GI:8953405"
	/db_xref="GOA:Q9LFU1"
	/db_xref="UniProt/TREMBL:Q9LFU1"
	/translation="WCGILAVLGVNDSQAKRSRIIELSRRLRHGPDWSGLHVCYEDC YLAHERLAIVDPTSGDQDPLYNEDKTIATVNGEYNNKALRENLSKHQFPGSDCEVI AHLYEEGEEFVMDGFAVLIDTRDKSFIARDAIGITPLYIGWJGJGVSWFASE MKALSDCECEFCFPGHIYSKQGLRRWYNPPWSPVSTPYDPLVLRNTPEKAV IKRLMTDVPFGLVLLSGDLSLVALEKSEAAACQWGSKLHTFCIGKSGPDLK AGREVADYLGTRHRLHFTVQGDIAIEVIYHVETVDTTIRASTPFLMSRKIKSL

exon	complement (17729. .17941) /gene="F18D22_10" /number=1
intron	complement (17942. .18037) /gene="F18D22_10" /number=1
exon	complement (18038. .18145) /gene="F18D22_10" /number=2
intron	complement (18146. .18236) /gene="F18D22_10" /number=2
exon	complement (18237. .18323) /gene="F18D22_10" /number=3
intron	complement (18324. .18408) /gene="F18D22_10" /number=3
exon	complement (18409. .18489) /gene="F18D22_10" /number=4
intron	complement (18490. .18698) /gene="F18D22_10" /number=4
exon	complement (18699. .18833) /gene="F18D22_10" /number=5
intron	complement (18834. .18916) /gene="F18D22_10" /number=5
exon	complement (18917. .19138) /gene="F18D22_10" /number=6
intron	complement (19139. .19222) /gene="F18D22_10" /number=6
exon	complement (19223. .19303) /gene="F18D22_10" /number=7
intron	complement (19304. .19388) /gene="F18D22_10" /number=7
exon	complement (19389. .19550) /gene="F18D22_10" /number=8
intron	complement (19551. .19644) /gene="F18D22_10" /number=8
exon	complement (19645. .19739) /gene="F18D22_10" /number=9
intron	complement (19740. .19923) /gene="F18D22_10" /number=9
exon	complement (19924. .20019) /gene="F18D22_10" /number=10
intron	complement (20020. .20121) /gene="F18D22_10" /number=10
exon	complement (20122. .20263) /gene="F18D22_10" /number=11
intron	complement (20264. .20452) /gene="F18D22_10" /number=11
exon	complement (20453. .20548) /gene="F18D22_10"

```

Db 44538 GTACAAAGCTCAAGCCCTCCTTGAATCAGCCGCCACCGGTTACTGCTGTCAAACTCGC 44597
|||||
QY 181 CGCTCCGGGATCGACCCAGCACCAATCCAGGCCGTTGAAACCGTCGACGAACTATCCAA 240
Db 44598 CGCTCCGGGATCGACCCAGCACCAATCCAGGCCGTTGAAACCGTCGACGAACTATCCAA 44657
|||||
QY 241 ACCGTTGATCGAAACCCCTCAAATCTCTTCCAGAGAACGAAAGAGTAATTCTGTTGGATT 300
Db 44658 ACCGTTGATCGAAACCCCTCAAATCTCTTCCAGAGAACGAAAGAGTAATTCTGTTGGATT 44717
|||||
QY 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGCGACATATTTCCGGCGAAGATTAAAGGT 360
Db 44718 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGCGACATATTTCCGGCGAAGATTAAAGGT 44777
|||||
QY 361 TCTTGTGTTCTCAAGCGCTTCTTGCCCGACACCAACCCAGTCGCTTCTCAGTTCCTGGA 420
Db 44778 TCTTGTGTTCTCAAGCGCTTCTTGCCCGACACCAACCCAGTCGCTTCTCAGTTCCTGGA 44837
|||||
QY 421 CAAG 424
|||||
Db 44838 CAAG 44841
|||||

RESULT 6
AY202287 856 bp DNA linear SYN 27-MAY-2003
LOCUS Arabidopsis thaliana sequence flanking Ds3 end of Ds-Trap insertion
DEFINITION from line GT6478.
ACCESSION AY202287
VERSION AY202287.1 GI:27898241
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 856)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE May B.P., Simorowski J., Arroyo J.-M., Vaughn, M.W., Shen, R.,
AUTHORS McCombie, W.R. and Martienssen, R.A.
COMMENT Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 856)
AUTHORS May B.P., Simorowski J., Arroyo J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
COMMENT http://genetraps.cshl.org.
FEATURES
source
location/Qualifiers
1..856
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT6478.Ds3.05.08.98.b.888"
/ecotype="Landsberg erecta"
/transgenic
/notes="transgenic line GT6478"
1..34
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
repeat_region <1..34
/transposon="transposon Ds"

ORIGIN
Query Match 15.4%; Score 140; DB 12; Length 856;
Best Local Similarity 99.0%; Pred. No. 5.7e-67;
Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 555 AGGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTAT 614
|||||

```

```

Db 336 AGGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTAT 395
|||||
QY 615 CAAAGAAAGAAAGTTTAGCGAGGAGGATATGGTTCGGTGCACACGAGTTTTCGTAATGA 674
Db 396 CAAAGAAAGAAAGTTTAGCGAGGAGGATATGGTTCGGTGCACACGAGTTTTCGTAATGA 455
|||||
QY 675 GTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTGTTGGATGATTAATTTTCAACG 734
Db 456 GTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTGTTGGATGATTAATTTTCAACG 515
|||||
QY 735 TCTCGAAGTCTACGAGATCGATGGCGGAGATCACATGCTGCTCTCAACACCCCAA 794
Db 516 TCTCGAAGTCTACGAGATCGATGGCGGAGATCACATGCTGCTCTCAACACCCCAA 575
|||||
QY 795 AACTCTTTTCACTCTCTCTGCTATTGCCCACGATTATATGTAATAATCTTAA 847
Db 576 AACTCTTTGACTCTCTCTGCTATTGCCCACGATTATATGTAATAATCTTAA 628
|||||

RESULT 7
BX321957 208 bp DNA linear STS 10-JUN-2003
LOCUS Arabidopsis thaliana transposon insertion STS SM_3.35356, sequence
DEFINITION tagged site.
ACCESSION BX321957
VERSION BX321957.1 GI:29420086
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1
AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 208)
AUTHORS Clarke, J.H.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT AT denotes an activation tag dissociation transposon within a
single line, ET an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. _3 denotes a sequence derived from the 3' end of the
transposon. _5 denotes a sequence derived from the 5' end of the
transposon. BBSRC GARNET, ARIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N122067.
FEATURES
source
location/Qualifiers
1..208
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL360334"
/notes="Derived from superpool 26.07 NASC code Unknown"
1..208
/standard_name="SM_3.35356"

STIS
1..208
/standard_name="SM_3.35356"

ORIGIN
Query Match 14.6%; Score 133; DB 11; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.8e-63;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 780 TCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATGTAAT 839
Db 1 TCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATGTAAT 60
|||||
QY 840 AATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTCCG 899
Db 61 AATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTCCG 120
|||||

```





Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hodges, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jarkson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowic, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajis, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczak, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 238185)  
Worley, K.C.

Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 238185)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 13, 2002 this sequence version replaced gi:23603976.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
-----  
Center project name: GMGO  
Center clone name: CH230-33309  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 226513 bases at least Q40  
Consensus quality: 228567 bases at least Q30  
Consensus quality: 229730 bases at least Q20

Estimated insert size: 234005; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 231322: contig of 231322 bp in length  
\* 231323 231422: gap of unknown length  
\* 231423 235627: contig of 4205 bp in length  
\* 235628 235727: gap of unknown length  
\* 235728 236737: contig of 1010 bp in length  
\* 236738 236837: gap of unknown length  
\* 236838 238185: contig of 1348 bp in length.

FEATURES  
Location/Qualifiers  
1..238185  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-33309"  
59308..60745  
/note="wgs\_contig"  
231423..232441  
/note="wgs\_contig"

misc\_feature  
misc\_feature  
ORIGIN

Query Match 2.5%; Score 23; DB 2; Length 238185;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 452 GATTGTGAGTTTTCATCTCATGA 474  
|||||  
DB 135161 GATTGTGAGTTTTCATCTCATGA 135139  
|||||

RESULT 11  
AP003765  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,  
BAC clone:OJ1003\_C06.  
ACCESSION  
AP003765 GI:50509923  
VERSION  
AP003765  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
1  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
clone:OJ1003\_C06  
Published Only in Database (2001)  
2 (bases 1 to 143168)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (20-JUN-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Jul 22, 2004 this sequence version replaced gi:34393713.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH  
(http://www.softberry.com/), GeneMark.hmm  
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM  
(http://www.tigr.org/tdb/glimmerm/glmr\_form.html), RiceHMM

(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OJ1003\_C06 clone has an overlap with P0524G08 clone (DDBJ: AP004671) at 5' end and an overlap with OJ1458\_B07 (DDBJ: AP003832) at 3' end. The sequence was generated by combining MonSanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/Genomeseq.html.

FEATURES  
source

1. .143168  
/location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="7"  
/clone="OJ1003\_C06"  
complement (1942..4991)  
/gene="OJ1003\_C06.101"  
complement (join(<1942..2007,2095..2154,2618..2658,3821..3854,3935..3985,4926..>4991))  
/gene="OJ1003\_C06.101"  
/note="start and end point are not identified"  
complement (join(1942..2007,2095..2154,2618..2658,3821..3854,3935..3985,4926..4991))  
/gene="OJ1003\_C06.101"  
/note="contains EST(s): AU030351 (E50932), AU030350 (E50932)"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="BAD30245.1"  
/db\_xref="GI:50509924"  
/translation="MDERFRMLFLPVVRSDYCAESGSSKGSTQCSQAQETRGNKSSSDQFMKKLKWAEKVGATKAEFLCKTFEAKKKLVTKELNLDAAQFLNAYE KGS"  
complement (6337..6991)  
/gene="OJ1003\_C06.102"  
complement (join(<6337..6615,6764..>6991))  
/gene="OJ1003\_C06.102"  
/note="start and end point are not identified"  
complement (join(6337..6615,6764..6991))  
/gene="OJ1003\_C06.102"  
/note="predicted by FGENESH etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30246.1"  
/db\_xref="GI:50509925"  
/translation="WADRRPAYNTTKARLQKRCPHGGHDKDAIAHLRKMWFPPKKE NSPRKGYRPOEGYDTRCSRCPGDHNPAASSLHPTAPPLRAGPGHLRPATR RPPOVRPKENPDWTVSLRPPELPLLTSSRLNPTETPHKEGFEFKKGADWQOQGRRTAASSHEH"  
complement (8772..9955)

mrna  
complement (join(<8772..8959,9611..9796,9898..>9955))  
/gene="OJ1003\_C06.103"  
/note="start and end point are not identified"  
complement (join(8772..8959,9611..9796,9898..9955))  
/gene="OJ1003\_C06.103"  
/note="predicted by GeneMark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30247.1"  
/db\_xref="GI:50509926"  
/translation="MEQTGSSGSGQGLAKGLHDANVAYVMGKYDASDEKAVLAVDM ANKRLQIVSVYDAERIYDDPDYSYQTSISQYFTTAAGVKGKLNKRLPKFHAQYPHKRL GETLSRDNPMDLHEPQLQDIFGLGTDETDSDNPMDLF"  
complement (10031..10462)  
/gene="OJ1003\_C06.104"  
complement (<10031..>10462)  
/gene="OJ1003\_C06.104"  
/note="start and end point are not identified"  
complement (10031..10462)  
/gene="OJ1003\_C06.104"  
/note="predicted by GeneMark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30248.1"  
/db\_xref="GI:50509927"  
/translation="MATVFPYRSRTEHLTLRPPCHGGDPLDAHQPLPWILLDVRYAI ADRENSTTAADLGNHLYRPTAAGLILCLVHLTKVITLDEGQGEVAFIDLKRGI NPHLQARPMLSWSKPATAARHOLSNELOVLIALPGHCHR"  
complement (12401..20516)  
/gene="OJ1003\_C06.105"  
complement (join(<12401..12654,12751..12825,12909..12988,13056..13136,13227..13483,13578..13806,13889..14043,14125..14244,14351..14429,14514..14650,14780..14882,14973..15246,15337..15487,18354..18509,18621..18711,19872..19911,20211..>20516))  
/gene="OJ1003\_C06.105"  
/note="supported by full-length cDNA(s): AK100505"  
complement (12401..20516)  
/gene="OJ1003\_C06.105"  
/note="contains full-length cDNA(s): AK100505  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
complement (20915..23256)  
/gene="OJ1003\_C06.106"  
complement (join(20915..21640,22598..22819,22922..23256))  
/gene="OJ1003\_C06.106"  
/note="supported by full-length cDNA(s): AK058442"  
complement (join(20985..21640,22598..22819,22922..23219))  
/gene="OJ1003\_C06.106"  
/note="contains full-length cDNA(s): AK058442"  
/codon\_start=1  
/product="lipase-like protein"  
/protein\_id="BAD30249.1"  
/db\_xref="GI:50509928"  
/translation="MATRPARLSAPAVAILVAALCAATASAAVNTTTATAPSGG CYSHLFTGNSLIDGNFIHYSTPGPVARSFYGETFRRTGRWSDGRLIVDFIVER LGFPYWPYLAGSKRDFRYGANFASGTALNQLLKKGLSVAGITPYSLAVQVGW PKYVLAMLASTEQERKEAMARSVEFMGDFGNDYLLHPLFQNTLWVRLPVVRYI AGAVEELVGLGATTYYVGPFLPCVPRLLFLPRDGGAGDRDPATGCLRGDLGAL HALLRLRLAELRAHPGVTTAYADYGEVMELVSNPAASGFDALTACCAGGGPYNG NHTVHCSDPGATQCAQDFSRISDGLHMTAVYRIMARGVLDPADPPIMSRCHGY" 24414..26326  
/gene="OJ1003\_C06.107"  
join(<24414..24473,24569..24666,25426..25525,25698..25880,26204..>26326)  
/gene="OJ1003\_C06.107"  
/note="start and end point are not identified"  
join (24414..24473,24569..24666,25426..25525,25698..25880,26204..26326)  
/gene="OJ1003\_C06.107"  
/note="contains EST(s): C74320 (E30875), AU088598 (C53863), AU166063 (E30875), C97298

```

(C53863)"
/codon_start=1
/product="transport protein particle component Bet3-like
protein"
/protein_id="BAD30250.1"
/db_xref="GI:50509929"
/translation="MPPLTPGKSGDALFGSVVERVNAELFTLYGAIVRQLLFDLBEVE
EVNKQIDMGVYIGTRLDVDFLAKNSVRCVDFKFTADVIAGLGFPMFLGVATVTNW
DAEGTSCVLENDPLVDVFLPDTCQGLQYCNVLSGVIRGALEVMVKTEVTWVRDM
LRGDDAYENRVKLTQVPEEYPKDD"
27936..31529
/gene="OJ1003_C06.108"
/notes="OJ1003_C06.108"
/translation="WASTSRGSGEDDDDEPYLLGFIIVANVLOYOGGRAITRES
VGLVREPHNPHANAIRVDNARGEKIGHIRRAAALAPLLDAGHVAAHGIVPKPAS
KRLYSLPQCQHLFARPPHAALAAASGIDLIIHVDHPEFALSESAIVEQQTKRSR
2.4%; Score 22; DB 8; Length 143168;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 CTCGCCGCTCCGGGATGACC 197
Db 28317 CTCGCCGCTCCGGGATGACC 28338

RESULT 12
AP004671
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
PAC clone:P0524G08.
ACCESSION
AP004671
VERSION
AP004671.2 GI:22002483
KEYWORDS
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
Clone:P0524G08
Published Only in Database (2002)
2 (bases 1 to 146690)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (23-JAN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Jul 29, 2002 this sequence version replaced gi:18307749.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://opal.biology.gatech.edu/glimmer/glmr form.html), RiceHMM
(http://rgp.dna.affrc.go.jp/RiceHMM/), _SplicePredictor
(http://biinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DDBJ. Protein homologies of the coding

```

```

regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-', and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0524G08 clone has an overlap with OJ1340_C08
(DDBJ: AP005292) clone at 5' end and with OJ1003_C06 (DDBJ:
AP003765) at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1..146690
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="P0524G08"
1573..3882
/gene="P0524G08.101"
join(1573..1933,3403..3841)
/notes="supported by full-length cDNA (s): AK059275"
join(1590..1933,3403..3882)
/gene="P0524G08.101"
/notes="supported by full-length cDNA (s): AK070897"
join(1721..1933,3403..3585)
/gene="P0524G08.101"
/notes="contains Set's; AU162548 (R3914), AU032339 (R3914)
contains full-length cDNA (s): AK059275, AK070897
similar to Arabidopsis thaliana chromosome 3, At3g12260"
/codon_start=1
/product="unknown protein"
/protein_id="BAC16465.1"
/db_xref="GI:23237891"
/translation="MAFTMRVAVKVPVNSASLEEARHVRVDFFRQACRAIPSIIMEIYNL
DDVVTSQLRSTIAKEIRKNQGVNPKVIDMLFKGMEELGNITEHAKQRHVIQYV
VGQKGLVQDMKQDQSSDFLKFKFTSNYS"
4758..9355
/gene="P0524G08.102"
join(4758..5101,6083..6734,6851..6939,7398..7462,
7571..7626,7711..7831,7920..7994,8293..8346,8434..8528,
8608..8782,8897..9355)
/gene="P0524G08.102"
/notes="supported by full-length cDNA (s): AK070161"
join(5040..5101,6083..6734,6851..6939,7398..7462,
7571..7626,7711..7831,7920..7994,8293..8346,8434..8528,
8608..8782,8897..9141)
/gene="P0524G08.102"
/notes="contains full-length cDNA (s): AK070161"
/codon_start=1
/product="putative xylose kinase"
/protein_id="BAC16466.1"
/db_xref="GI:23237892"
/translation="WVGRSSLPEGSFLGFDSTQSLKATVLNNELIIVASDIVNFD
DLPHYKTEGVYRPADGGHIFPSIMVVEAFELLLEKPKKINFKVAVLSGSGQOH
GSVTTWKGGSHVLSLSDPTKLSLQDKDAFTMSDPSIMWDSSTTKQCREISVUGGAL
ELSKLTGSRAVERFTGQRIKYRTVPQVYDDTERISLVSSFMASILVGNVAFIDETD
CAGNMLADINQRTWSTKTVLEATPGLEKGLKAPAVAGRIAPFYFVERIQDNCL
VIOWSGDNPNLSLAGLTNTPGDLAISLGTSDTVFGITAEAKPSLEGHVFPNPEPDGY
MVMLCYKNGSLTREDVRNSCAEKSWDVFNSTLEKTPPLNGKGLGFYFKHDLPLPLV

```

## FEATURES

source

GFHRVIVENLNDVTSNNLVERVEFPDPSEVRATIEIQQLLSMRCHAEFQWNPPKR  
IATGASNERILHSIAQIFCCPVFTVQRPDSASLGAALRAAGNLCNBSGFPVPS  
CMYQGNLEKTSLGAKLAVSTGEGVELLEKYTVLMKRMEIERRLVEKIGRA"  
complement(10645..14070)  
/gene="P0524G08.103"  
complement(join(<10645..10870,11911..12051,12128..12186,  
12255..12561,13086..13167,13244..13377,13462..13546,  
13626..13738,13823..13902,14011..>14070))  
/gene="P0524G08.103"  
/note="start and end point are not identified"  
complement(join(10645..10870,11911..12051,12128..12186,  
12255..12561,13086..13167,13244..13377,13462..13546,  
13626..13738,13823..13902,14011..14070))  
/gene="P0524G08.103"  
/note="predicted by GeneMark.hmm etc.  
similar to Oryza sativa chromosome 8, OJ1081\_B12.12"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16467.1"  
/db\_xref="GI:23237893"  
/translation="MASASTPTIFSIDGVILDEBTMMQFTIPKKLENEFRPSLTLC  
VITKVHLEGRSGIIPNYGEYFCDFDQIPSKSRISKPLIGVISHVGPDYAGKTSSK  
KNRKLIRKDEQEGIVLWGEYGSFDEAFVLOKSTDHKIWAIVLGLTAGTVLTKT  
EATSLATQIVPDSITELIVEQSRYKPLATIKDELGTLDAAVFNVAEDLVEYNATQ  
ATQNKIDATEHAIALDTAIGKTRLFHAMTKYSSHFTINYVLKSPVENNTISLM  
LPLENTKVESATQKATDEGLTMEHYSKDQHPPTPPSLQPPETTLDDNTVNQI  
IPSAKRALQFEKELIDQPRNHEVQLIHTNYLQAVSKIASKSTAISITLQKGSSTD  
LSKLQTPQAHVVMYMKPLPHMHIILDGENVKNEI"  
complement(14288..18600)  
/gene="P0524G08.104"  
complement(14288..18600)  
/gene="P0524G08.104"  
/note="probably inactive due to including frameshift(s) in  
CDS  
pseudogene,helicase"  
/pseudo  
25168..27249  
/gene="P0524G08.105"  
join(<25168..25348,26073..26110,26865..27006,  
27095..>27249)  
/gene="P0524G08.105"  
/note="start and end point are not identified"  
join(25168..25348,26073..26110,26865..27006,27095..27249)  
/gene="P0524G08.105"  
/note="predicted by GENSCAN etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16468.1"  
/db\_xref="GI:23237894"  
/translation="MSRNDKSRAAVDRIKTAALSAAKGLNRSQKHAATSAAARNVPM  
GRKKDQAGRGRRLIILAAILVLEIHEAPLQFAGRAQANFAVHQLFVATAPSV  
GGIDGMGMQSCAAHMTASRSGFRAGRLIILADDSQDEBTSELSSTSSAWILC  
PWLLSPWIDD"  
30279..30467  
/gene="P0524G08.106"  
/note="start and end point are not identified"  
30279..30467  
/gene="P0524G08.106"  
/note="predicted by GlimmerM etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16469.1"  
/db\_xref="GI:23237895"  
/translation="MHGPAEAAAGRSIDGWLWTASLIVLALLAALMYQPYAPWPV  
QPCNCKPACVRQRSMTW"  
35625..36011  
/gene="P0524G08.107"  
35625..36011  
/gene="P0524G08.107"  
/note="supported by full-length cDNA(s) : AK107877"  
35625..36011

/gene="P0524G08.107"  
/note="contains full-length cDNA(s) : AK107877  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
join(37684..37833,39835..40872)  
/gene="P0524G08.108-2"  
join(37684..37833,39835..40872)  
/gene="P0524G08.108-2"  
/note="supported by full-length cDNA(s) : AK102159"  
join(37687..38060,39726..40877)  
/gene="P0524G08.108-1"  
join(37687..38060,39726..40877)  
/gene="P0524G08.108-1"  
/note="supported by full-length cDNA(s) : AK063952"  
join(37782..38060,39726..40379)  
/gene="P0524G08.108-2"  
/note="contains full-length cDNA(s) : AK063952,AK099354"  
  
Query Match 2.4%; Score 22; DB 8; Length 146690;  
Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0;  
  
Qy 176 CTCGCGCCTCCGGATCGACC 197  
|||||  
Db 99064 CTCGCGCCTCCGGATCGACC 99085  
  
RESULT 13  
AC010196 163246 bp DNA linear PRI 28-SEP-2002  
LOCUS Homo sapiens 12 BAC RP11-202H2 (Roswell Park Cancer Institute Human  
DEFINITION BAC Library) complete sequence.  
AC010196  
AC010196.5 GI:6139072  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 163246)  
AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,  
Bodota,B., Bock,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,  
Burkert,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,  
David,R., Delgado,O., Deshazo,D., Ding,Y., Donah-Rashid,N.,  
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,  
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,  
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,  
Holloway,C., Hosak,H., Issar,A., Jackson,L.E., Jackson,L., Jia,Y.,  
Jones,M., Kelly,S., Kneitz,S., Kondejewski,N., Kong,Y., Kovar,C.,  
Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,  
Logan,O., Lu,J., Lucier,R., Marondel,I., Martin,R., Martinez,C.,  
McLeod,M.P., Mei,G., Merscher,S., Miller,A., Montgomery,K.T.,  
Morgan,M., Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N.,  
Nguyen,S., Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L.,  
Pu,L.L., Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J.,  
Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M., Sparks,A.,  
Stamps,A., Sucgang,R., Tabor,P., Taylor,T., Vasquez,L., Vinson,R.,  
Vo,Q., Wahbah,M., Watlington,S., Weinstein,G., Weinstock,I.R.,  
Williamson,A., Worley,K., Wren,J., Wrensford,G., Zhang,A.M.,  
Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 163246)  
Worley,K.C.  
Direct Submission  
Submitted (15-SEP-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 163246)  
Worley,K.C.  
Direct Submission  
Submitted (29-OCT-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 163246)  
Worley, K.C.  
Direct Submission  
Submitted (31-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 163246)  
Worley, K.C.  
Direct Submission  
Submitted (28-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Oct 29, 1999 this sequence version replaced gi:6087859.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----  
Contig length: 80737  
Phrap values in estimate: 80641  
Average error rate (BCM-Phrap estimate): 8.31934e-05  
Fraction of Phrap values less than 40 : 0.02247  
Number of consensus changing edits: 0  
Number of N's in consensus : 3

----- Consensus changing edits -----  
Position Original+Context Edited+Context

----- Distribution of Quality < 40 Bases -----

1000	*
900	*
800	*
700	*

bases 600 \*  
500 \*  
400 \*  
300 \*  
200 \*  
100 \*  
0 \* \* \* \* \*  
-----  
5 10 15 20 25 30 35 40  
Phrap Value Range

Version: 1.01 gxf.  
Location/Qualifiers  
1. 163246  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-202H2"  
complement(1..2001)  
/note="overlaps bases 1..2001 of clone AC128681"  
/function="clone overlap"  
5715..5859  
repeat\_region  
/rpt\_family="MLT1E1"  
5860..6155  
repeat\_region  
/rpt\_family="AluJb"  
6156..8222  
repeat\_region  
/rpt\_family="MLT1E1"  
6223..6738  
repeat\_region  
/rpt\_family="MLT1D"  
6739..6817  
repeat\_region  
/rpt\_family="MLT1E1"  
complement(7107..7419)  
repeat\_region  
/rpt\_family="L1MB1"  
complement(7451..7662)  
repeat\_region  
/rpt\_family="L1MEC"  
7814..7841  
repeat\_region  
/rpt\_family="(CAAAA)n"  
complement(7842..8283)  
repeat\_region  
/rpt\_family="L1MA6"  
8281..9425  
repeat\_region  
/rpt\_family="L1MA6"  
9493..9514  
repeat\_region  
/rpt\_family="AT\_rich"  
9577..9809  
repeat\_region  
/rpt\_family="MIR"  
9812..9834  
repeat\_region  
/rpt\_family="AT\_rich"  
10808..10845  
repeat\_region  
/rpt\_family="AT\_rich"  
complement(10846..11182)  
repeat\_region  
/rpt\_family="L1PA7"  
11183..11231  
repeat\_region  
/rpt\_family="AT\_rich"  
11753..12148  
repeat\_region  
/rpt\_family="MLT1A2"  
12166..12244  
repeat\_region  
/rpt\_family="AT\_rich"  
12265..12554  
repeat\_region  
/rpt\_family="AluJb"  
12555..12669  
repeat\_region  
/rpt\_family="GA-rich"  
complement(12805..13216)  
repeat\_region  
/rpt\_family="MLT2E"  
complement(13709..14622)  
repeat\_region  
/rpt\_family="L1MB1"  
complement(14623..15117)  
repeat\_region  
/rpt\_family="MER9"  
complement(15118..15490)  
repeat\_region  
/rpt\_family="L1MB1"  
15534..15867  
repeat\_region

```
/rpt_family="LIM4"
17685..17724
/rpt_family="TTTGn"
20420..20441
/rpt_family="AT rich"
complement(21000..21182)
/rpt_family="MIR"
complement(21916..22025)
/rpt_family="L2"
23794..24307

Query Match      2.4%; Score 22; DB 9; Length 163246;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 278 GAAGAGGTAATCTGTTGGAT 299
|||||
Db 113646 GAAGAGGTAATCTGTTGGAT 113667

RESULT 14
AC087244/c
LOCUS
DEFINITION Homo sapiens 12 BAC RP11-444N1 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC087244
VERSION AC087244.17 GI:13470125
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 187111)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.J., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,X., Dinh,H.H., Douchwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Minzer,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,N., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokewko,S.,
Ogih,M., Okuwonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoshitari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Umani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
```

Zorrilla,S., Kucherlapati,R. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (20-DEC-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (28-MAR-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (02-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
6 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (02-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
7 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (03-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Mar 28, 2001 this sequence version replaced gi:13443049.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:  
STGs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base

quality are listed below. Description of the metrics can be found at URL:  
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

```
FEATURES
  source            Location/Qualifiers
    1..187111
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="12"
      /clone="RP11-444N1"
      complement(3..238)
      /rpt_family="L1MB8"
      complement(230..371)
      /rpt_family="L1MB8"
      complement(2548..2861)
      /rpt_family="L1MB4"
      3075..3176
      /rpt_family="L2"
      3654..3753
      /rpt_family="L2"
      4503..4734
      /rpt_family="MIR"
      4950..4997
      /rpt_family="(CAA)n"
      5275..5340
      /rpt_family="MIR"
      complement(6876..7173)
      /rpt_family="AluJo"
      8248..8280
      /rpt_family="L1PA13"
      8281..8678
      /rpt_family="L1PA13"
      8693..8922
      /rpt_family="L1MB6"
      8923..9222
      /rpt_family="AluSg"
      9223..9592
      /rpt_family="L1MB6"
      9782..9853
      /rpt_family="L2"
      9978..10106
      /rpt_family="MER115"
      complement(10107..10750)
      /rpt_family="L1MB3"
      10751..10882
      /rpt_family="MER115"
      11732..11798
      /rpt_family="L1MA4"
      11805..11825
      /rpt_family="AT_rich"
      12087..12109
      /rpt_family="AT_rich"
      12452..12628
      /rpt_family="L1MA9"
      13629..12670
      /rpt_family="(CA)n"
      12671..12894
      /rpt_family="L1MA9"
      12895..12922
      /rpt_family="AT_rich"
      complement(13539..13588)
      /rpt_family="L2"
      complement(15205..15428)
      /rpt_family="MIR"
      15883..15903
      /rpt_family="(T)n"
      complement(15904..16035)
      /rpt_family="FLAM_C"
      18104..18127
      /rpt_family="AT_rich"
      18531..18634

repeat_region      2.4%; Score 22; DB 9; Length 187111;
                    Best Local Similarity 100.0%; Pred. No. 3.7;
                    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 AGGAACATCACTTCGTGTAG 89
    |||||
Db 80631 AGGAACATCACTTCGTGTAG 80610

RESULT 15
AC025128/c
LOCUS              31740 bp      DNA      linear      HTG 13-JUL-2000
DEFINITION         Homo sapiens clone RP11-307E16, LOW-PASS SEQUENCE SAMPLING.
AC025128
AC025128.1 GI:7158939
VERSION            HTG; HTGS PHASE0.
KEYWORDS
SOURCE            Homo sapiens (human)
ORGANISM           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 31740)
Homo sapiens, clone RP11-307E16
Unpublished
2 (bases 1 to 31740)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,G.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7794
Center Clone name: 307_E_16
-----
* NOTE: This record contains 36 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
```

\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 799: contig of 799 bp in length  
800 899: gap of 100 bp  
1693: contig of 794 bp in length  
1694 1793: gap of 100 bp  
1794 2579: contig of 786 bp in length  
2580 2679: gap of 100 bp  
2680 3436: contig of 757 bp in length  
3437 3536: gap of 100 bp  
3537 4294: contig of 758 bp in length  
4295 4394: gap of 100 bp  
4395 5145: contig of 751 bp in length  
5146 5245: gap of 100 bp  
5246 6045: contig of 800 bp in length  
6046 6145: gap of 100 bp  
6146 6962: contig of 817 bp in length  
6963 7062: gap of 100 bp  
7063 7843: contig of 781 bp in length  
7844 7943: gap of 100 bp  
7944 8728: contig of 785 bp in length  
8729 9582: gap of 100 bp  
9583 9682: contig of 754 bp in length  
9683 10452: gap of 100 bp  
10453 10552: contig of 770 bp in length  
10553 11353: gap of 100 bp  
11354 11453: contig of 801 bp in length  
11454 12246: contig of 793 bp in length  
12247 12346: gap of 100 bp  
12347 13148: contig of 802 bp in length  
13149 13248: gap of 100 bp  
13249 14028: contig of 780 bp in length  
14029 14128: gap of 100 bp  
14129 14900: contig of 772 bp in length  
14901 15000: gap of 100 bp  
15001 15794: contig of 794 bp in length  
15795 15894: gap of 100 bp  
15895 16701: contig of 807 bp in length  
16702 16801: gap of 100 bp  
16802 17587: contig of 786 bp in length  
17588 17687: gap of 100 bp  
17688 18473: contig of 786 bp in length  
18474 18573: gap of 100 bp  
18574 19354: contig of 781 bp in length  
19355 19454: gap of 100 bp  
19455 20237: contig of 783 bp in length  
20238 20337: gap of 100 bp  
20338 21119: contig of 782 bp in length  
21120 21219: gap of 100 bp  
21220 21995: contig of 776 bp in length  
21996 22095: gap of 100 bp  
22096 22872: contig of 777 bp in length  
22873 22972: gap of 100 bp  
22973 23749: contig of 777 bp in length  
23750 23849: gap of 100 bp  
23850 24643: contig of 794 bp in length  
24644 24743: gap of 100 bp  
24744 25561: contig of 818 bp in length  
25562 25661: gap of 100 bp  
25662 26424: contig of 763 bp in length  
26425 26524: gap of 100 bp  
26525 27317: contig of 793 bp in length  
27318 27417: gap of 100 bp  
27418 28208: contig of 791 bp in length  
28209 28308: gap of 100 bp  
28309 29105: contig of 797 bp in length  
29106 29205: gap of 100 bp  
29206 29995: contig of 790 bp in length  
29996 30095: gap of 100 bp  
30096 30893: contig of 798 bp in length  
30894 30993: gap of 100 bp

\* 30994 31740: contig of 747 bp in length.

FEATURES  
source  
1. .31740  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-307E16"  
/clone\_lib="RPC1-II Human Male BAC"

ORIGIN

Query Match 2.3%; Score 21; DB 2; Length 31740;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 AGGCAAGGGTCATTTTCACA 604  
|||  
Db 15333 AGGCAAGGGTCATTTTCACA 15313

Search completed: September 12, 2005, 15:37:32  
Job time : 3966.19 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 20:27:23 ; Search time 183 Seconds  
(without alignments)

2324.765 Million cell updates/sec

Title: US-10-780-002-2

Perfect score: 1368

Sequence: 1 MKBGFHVLVHGACHGWSW.....MLCEPQKLCASLLETAHKYN 260

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DSV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US10780002/runat\_11092005\_104748\_26625/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002@cgn 1 1 105 @runat\_11092005\_104748\_26625 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfilee1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	585	42.8	1078	3	US-08-981-256A-4
2	141	10.3	840	4	US-09-328-352-1832
3	127	9.3	795	3	US-09-134-001C-2755
4	126	9.2	807	4	US-10-107-096-1
5	123	9.0	4331	4	US-09-710-279-3711
6	120.5	8.8	2101	1	US-08-106-761-1
7	117	8.6	1107	4	US-09-502-540-4129
8	117	8.6	9564	4	US-09-902-540-1026
9	114	8.3	8654	1	US-08-920-812-6
10	114	8.3	8654	1	US-08-920-827-6
11	114	8.3	8654	1	US-08-921-177-6
12	114	8.3	8654	1	US-08-362-577C-6

13	114	8.3	8654	2	US-08-920-828-6	Sequence 6, Appli
14	113	8.3	657	4	US-09-252-991A-9351	Sequence 9351, Ap
15	113	8.3	855	4	US-09-252-991A-9422	Sequence 9422, Ap
16	113	8.3	1101	4	US-09-252-991A-9455	Sequence 9455, Ap
17	109	8.0	855	4	US-09-328-352-1853	Sequence 1853, Ap
18	108	7.9	1116	4	US-09-602-777A-225	Sequence 225, App
19	105	7.7	891	4	US-09-902-540-6739	Sequence 6739, Ap
20	105	7.7	4009	4	US-09-902-540-541	Sequence 541, App
21	103	7.5	522	4	US-09-637-746-9	Sequence 9, Appli
22	103	7.5	1167	1	US-08-232-519-1	Sequence 1, Appli
23	103	7.5	1167	1	US-08-456-956-1	Sequence 1, Appli
24	102	7.5	834	4	US-09-252-991A-4372	Sequence 4372, Ap
25	100.5	7.3	873	4	US-09-248-796A-2815	Sequence 2815, Ap
26	99	7.2	1897	4	US-09-637-746-1	Sequence 1, Appli
27	99	7.2	2181	4	US-09-248-796A-3259	Sequence 3259, Ap
28	96.5	7.1	2772	4	US-09-540-236-1917	Sequence 1917, Ap
29	96.5	7.1	269223	4	US-09-596-002-41	Sequence 41, Appl
30	96	7.0	49272	1	US-08-614-770A-1	Sequence 1, Appli
31	95.5	7.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
32	95.5	7.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
33	94	6.9	885	3	US-09-424-349A-1	Sequence 1, Appli
34	93	6.8	522	4	US-09-637-746-7	Sequence 7, Appli
35	93	6.8	735	4	US-09-902-540-6017	Sequence 6017, Ap
36	93	6.8	1272	4	US-09-252-991A-14082	Sequence 14082, A
37	93	6.8	1329	4	US-09-252-991A-14036	Sequence 14036, A
38	93	6.8	1620	4	US-09-252-991A-14186	Sequence 14186, A
39	93	6.8	11029	4	US-09-902-540-1004	Sequence 1004, Ap
40	92	6.7	360	4	US-08-956-171E-1291	Sequence 1291, Ap
41	92	6.7	360	4	US-08-781-986A-1291	Sequence 1291, Ap
42	92	6.7	1796	1	US-07-816-283-11	Sequence 11, Appl
43	92	6.7	1796	1	US-08-417-103-11	Sequence 11, Appl
44	91	6.7	1316	4	US-09-461-325-91	Sequence 91, Appl
45	91	6.7	1316	4	US-10-012-542-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1  
US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:  
; APPLICANT: Mainhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981,256A  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob  
; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-OZ1112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:



```

Db      514  CAAGGCCAAATAATTATACATAGACCACAACTCGGGAAGAAAGTGCAATTGCA-----56
Qy      176  ValArgProSerLeuPheMetGluAspLeu-----SerLysAlaLysTyrPhe 192
Db      565  GTACGT---GCATTGCTTAATTATAAGATTTAACCTCGTTCATCATATAATGTGTCCATA 621
Qy      187  -----
Db      622  CCTACTTTTAATTGTCAATGGTAAGTATGACCCACTCATACAAAATAAAAGTCATTATGAT 681
Qy      193  ThrAspGluArgPheGlySerValIysArgValTyrIleValCysThrGluAspLysGly 212
Db      682  ATGGATCAATATTATGATCAAGTACAAATTTGTTATTT-----GATAATTCAGGA 732
Qy      213  -----IleProGluGluPheGlnArgTrpGlnIleAsp 223
Db      733  CATGCACCACATATCGAGGACACAGAAAATTCCTGAAACTCTACTTAGAT 783

RESULT 4
US-10-107-096-1
; Sequence 1, Application US/10107096
; Patent No. 6812380
; GENERAL INFORMATION:
; APPLICANT: Karlovsky, Petr
; APPLICANT: Gilliam, Jacob T.
; APPLICANT: Maddox, Joyce R.
; APPLICANT: Crane III, Edmund H.
; TITLE OF INVENTION: Compositions and Methods of Zearenone
; TITLE OF INVENTION: Detoxification
; FILE REFERENCE: 35718/245827
; CURRENT APPLICATION NUMBER: US/10/107,096
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/279,098
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/318,747
; PRIOR FILING DATE: 2001-09-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Gliocladium roseum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(807)
US-10-107-096-1

Alignment Scores:
Pred. No.: 2.05e-06 Length: 807
Score: 126.00 Matches: 60
Percent Similarity: 37.55% Conservative: 32
Best Local Similarity: 24.49% Mismatches: 121
Query Match: 9.21% Indels: 32
DB: Gaps: 10

US-10-780-002-2 (1-260) x US-10-107-096-1 (1-807)
Qy      27  LeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGly-----43
Db      130  CTGTGCGCGCGGCTACCGCGTGGCGCGCTCGATCTGCTGGACGACGCGACAGCACC 189
Qy      44  ThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db      190  GCAAACTTCAACCGCTACCGCGACGAGCGACCGCGCGACTAC-----CTT 237
Qy      64  GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db      238  ACAATTATTCAAACTCTGGGCGCAGAC---CGCGCGCTCTGGCAGGTGCCTCGATGTGC 294
Qy      84  GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db      295  GCTGCCGCGGCTTACATTGGCGGAGGAAAAACAACCCGACAAGATCCCGCGCTGGTCTT 354

```



```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1703
US-08-106-761-1

Alignment Scores:
Pred. No.: 4,91e-05 Length: 2101
Score: 120.50 Matches: 59
Percent Similarity: 38.16% Conservative: 28
Best Local Similarity: 25.88% Mismatches: 80
Query Match: 8.81% Indels: 61
DB: 1 Gaps: 11

US-10-780-002-2 (1-260) x US-08-106-761-1 (1-2101)
QY 6 HisPheVal-----LeuValHisGlyAlaCysHisGlyGly 17
Db 792 CATTTTGTGGAGCTGGCTGGCTGTGTGTGCTCTGCCATGGATTTCGCCGAGAGTTGG 851
QY 18 TrpSerTrpTyrLysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeu 37
Db 852 TATTTCTGTGGAGTACCAGATCCCTGTCTGTGGCCAGGCAGGTATCCGGGTCTCTAGTATG 911
QY 38 AspLeuAlaAlaSerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAsp 57
Db 912 GACATGAAGGCTATGGAGAGTCATCTGCTCCCTCCCGAA-----ATAGAGAA 959
QY 58 TyrThrLeu-----ProLeuMetGluLeuMetGluSerLeuSerAlaAsp 72
Db 960 TATTGCTGAAGTATTATGTAAGGAGATGTAACCTCTCTGGATAAACTGGGCGCTCTCT 1019
QY 73 GluLysValIleLeuValGlyHisSerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGlu 92
Db 1020 CAA---GCAGTGTTCATTGGCCATGACTGGGGTGGCATGCTGGTGTGG-----1064
QY 93 LysTyrProGlnLysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAspSerVal 112
Db 1065 -----TACATGGCTCTCTTC-----1079
QY 113 HisAsnSerSerPheValLeuGluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeu 132
Db 1080 -----TACCCCGAGAGAGTGGGGCGGTGGCCAGTTTG 1112
QY 133 AspThrGlnPheLeuProTyrGlySerProGluGluProLeuThrSerMetPheGly 152
Db 1113 AATACTCCCTTCATACAGCAATCCCAACATGCTCCCTTTTGAGAGATATCAAGCCAAAC 1172
QY 153 ProLysPheLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAla 172
Db 1173 CCAGTATTT-----GATTACAGCTCTACTTCCAAAGAACCCAGAGTGGCTGAG 1220
QY 173 SerSerLeuValArgProSerSerLeuPheMetGluAspLeuSerLysLysTyrPhe 192
Db 1221 GCTGAATCGAACAGAACCTGAGTCGGACTTTTCAAAAGCCTCTTCAGAGCA-----1271
QY 193 ThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysGly 212
Db 1272 AGCGATGAG-----AGTGTGTTTATCCATGTCATAAAGTCTGT---GAACGGGAGGA 1319
QY 213 Ile-----ProGluGlu 216
Db 1320 CTTTTTGTAAATAGCCAGAGAG 1343

RESULT 7
US-09-902-540-4129
; Sequence 4129, Application US/09902540
```

```
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4129
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4129

Alignment Scores:
Pred. No.: 5,01e-05 Length: 1107
Score: 117.00 Matches: 65
Percent Similarity: 41.73% Conservative: 41
Best Local Similarity: 25.59% Mismatches: 106
Query Match: 8.55% Indels: 42
DB: 4 Gaps: 14

US-10-780-002-2 (1-260) x US-09-902-540-4129 (1-1107)
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
Db 154 GACGGGAAGGCGGTGCTTTCATCCACGGCTGGGTTCGTACCTGAAGTCTTGGCGGGCG 213
QY 23 LysLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaLaser 42
Db 214 CAGTGGACGGCTTCCAGAACGAGGCTACCGCGTCATCGCGGTGAGCTGCCCGGTAC 273
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeu 62
Db 274 GGC-----AATCCGACAAAGCCCGCACCC-----TTCCTGACACCATGGAGGCC 318
QY 63 Met-----GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeu 77
Db 319 ATGGCGGACGGGTGCTGAGTTGGTGACGCGCTGGGCGCTGGAC---AAGCCCGTGTCTC 375
QY 78 ValGlyHisSerLeuGlyGlyMet---AsnLeuGlyLeuAlaMetGluLysTyrProGln 96
Db 376 GCGGTCTACTCATGGCGGGGCGACGCTGCTCTCTTCCGCATC---CGCTACCCCGGAG 432
QY 97 LysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSer 116
Db 433 TCGCTGACGGGGTGTGCTGCTGGCTGGCCC-----GCCGCG 468
QY 117 PheValLeuGluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPhe 136
Db 469 TTCGAGAAGTTTCAGCTGGCGGGAAGAGGTGTTTCGCGCGGCGCATGATGCTCCGAGTTC 528
QY 137 LeuProTyrGlySerProGluGluProLeu-----ThrSerMet 149
Db 529 ATCAAG---TCCGCGCGCGGAGGCCAGCATCTGGGGCAGCGTGGCGCACGCAACTTCATG 585
QY 150 PhePheGlyPro-----LysPheLeuAlaHisLysLeuTyrGlnLeuCys---SerPro 166
Db 586 CACTGGCGCGCGAGTTGGATGGCTGATTGAGGAGCGCGCTGGCTGGCGAAGTCGCC 645
QY 167 GluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGluAspLeu 186
Db 646 GAGTTCGACGCTACGCTACGCCAACGTGGCAGC-----GTGCGAGGCGCTC 693
QY 187 SerLysAlaLysTyrPheThrAspGluArgPheGly---SerValLysArgValTyrIle 205
Db 694 TCCCAACACGACTTCGTGGCGGACAACTGCACCGCGTCACCGTCCCGCCCGCTCATCATC 753
```

QY 206 ValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsnIle 225  
|||:||||:|||||  
Db 754 TACGGCAGCGATGACGGCTCATCCCGACGCCCTTCTCCACGGC----- 798  
QY 226 GlyValThrGluAlaIleGluIleLysGlyAlaAspHisMet 239  
|||:||||:|||||  
Db 799 GGAGAGACGGCGGCATCATGAGTACGGCGGTCCACATT 840

## RESULT 8

US-09-902-540-1026  
; Sequence 1026, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1026  
; LENGTH: 9564  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-1026

Alignment Scores:  
Pred. No.: 0.00158 Length: 9564  
Score: 117.00 Matches: 65  
Percent Similarity: 41.73% Conservative: 41  
Best Local Similarity: 25.59% Mismatches: 106  
Query Match: 8.55% Indels: 42  
DB: 4 Gaps: 14

US-10-780-002-2 (1-260) x US-09-902-540-1026 (1-9564)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTrpLys 22  
:::|||||:|||||  
Db 1676 GAGCGAAGCGCGTCTCTTCATCCAGCGCCTGGGTCTGACTGAAGTTCTGGCGGCG 1735  
QY 23 LeuLysProLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSer 42  
:::|||||:|||||  
Db 1736 CAGCTGGAGCGCTTCAGAGAAGAGGCTACCGCGTCATCGCGTGGACCTGCCGCTAC 1795  
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeu 62  
|||:|||||:|||||  
Db 1796 GGC-----AAGTCCGACAGCCCGCACCC-----TTCCCGTACACCATGGAGGCC 1840  
QY 63 Met-----GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeu 77  
|||:|||||:|||||  
Db 1841 ATGGCGGACGGTGTGAGTTGTGGACCGCCTGGGCTGGAC---AAGCCCGTGTCT 1897  
QY 78 ValGlyHisSerLeuGlyMet---AsnLeuGlyLeuAlaMetGluLysTrpProGln 96  
|||:|||||:|||||  
Db 1898 GCCGGTCACTCCATGGCGGGCAGAGCTCGCTCTCTTCGCCATC---CGCTACCGGGAG 1954  
QY 97 LysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSer 116  
|||:|||||:|||||  
Db 1955 TCGCTGAGCGGTGGTGTGGTGGTGGCC-----GCCGGC 1990  
QY 117 PheValLeuGluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPhe 136  
|||:|||||:|||||  
Db 1991 TTCGAGAAGTTCACTGGCGGGAGAGAGGTGTTCCGCCGGCGATGAGCTCCGAGTTC 2050  
QY 137 LeuProTyrGlySerProGluGluProLeu-----ThrSerMet 149  
|||:|||||:|||||  
Db 2051 ATCAAG---TCGGCGCGGAGGCGCAGCATCTGGGGCAGCGTGGCGGCGCCCAACTTCATG 2107  
QY 150 PhePheGlyPro-----LysPheLeuAlaHisLysLeuTyrGlnLeuCys---SerPro 166

Db 2108 CACTGGCGCGGAGTGGAGTGGCTGATTGAGAGCGCGTGGCGTGGCGAAGTGGCCC 2167  
|||:|||||:|||||  
QY 167 GluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGluAspLeu 186  
|||:|||||:|||||  
Db 2168 GAGTTTCGAGCGCTACGCTACGCCAACAGTGGCAGC-----GTCCGAGGCGCTG 2215  
QY 187 SerLysAlaLysTyrPheThrAspGluArgPheGly---SerValLysArgValTyrIle 205  
|||:|||||:|||||  
Db 2216 TCCACACACGACTTCGTGCGCGCACACCTGCACCGCGTCACTGGTCCCGCATCATC 2275  
QY 206 ValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsnIle 225  
|||:|||||:|||||  
Db 2276 TACGGCAGGATGACCGGCTCATCCAGAGCCCTTCTCCACGGC----- 2320  
QY 226 GlyValThrGluAlaIleGluIleLysGlyAlaAspHisMet 239  
|||:|||||:|||||  
Db 2321 GGAGAGACGGCGGCATCATGAGTACGGCGGTCCACATT 2362

## RESULT 9

US-08-920-812-6  
; Sequence 6, Application US/08920812  
; Patent No. 5763188  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,812  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8654 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus epidermidis  
; STRAIN: Clinical Isolate SB-22  
US-08-920-812-6

Alignment Scores:  
Pred. No.: 0.0033 Length: 8654  
Score: 114.00 Matches: 62  
Percent Similarity: 40.31% Conservative: 42

```

Best Local Similarity: 24.03% Mismatches: 97
Query Match: 8.33% Indels: 57
DB: 1 Gaps: 12

US-10-780-002-2 (1-260) x US-08-920-812-6 (1-8654)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTrpSerTrpTrpLys 22
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 681 GAAGGCATCCCATCATTTTAATTCATGGA---TTAGATGGAACCTGGCAGGATTAAA 737

QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 738 GATTATAAATAAGAACTCAAGAGCAG---TATAGAGTAATTAATTAATGATGTCAGAGT 794

QY 42 SerGlyThrAspLeuArgThrLeuLeuGluLeuArgThrLeuLeuAspTrpThrLeuPro 61
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 795 CATGGAAATCTTCACGA-----ACAGAATCATATGAATTAAGAATCATGTTGAAGAT 848

QY 62 LeuMetGluLeuMetGluSerLeuAlaAspGluLysValIleLeuValGlyHisSer 81
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 849 TTAATGATTAATGGAGCATTAATATCGATTCGCACATATTTTA---GGACATGAT 905

QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTrpProGlnLysIleTrpAlaAla 101
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 906 ATGGGGGCATCATTCGAGTGAATTTACTGAAATAATCAATATAAAGTGATTACATTG 965

QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 966 ACAATTGTTTCGGCCAAAAGTGAAGACATTCGAAATGGTTTCAACAAATTAATGGTTGAT 1025

QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1026 TACCAAGAGAATTAGCAGGCTTTAATAATCTGAGGCAATGATTATT----- 1073

QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1074 -----TTATTTCTTAATTTTAAAGAGAAGATAAAGCAATGAATGGGTA 1121

QY 155 eLeuAlaHisLysLeuTrpGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLe 175
   : : : : :
   : : : : :
   : : : : :
   : : : : :
Db 1122 TCAAGCCAAAATTTATACAAATAGACCAACTCCGGAAGAAAGTGCAATTGCA----- 1173

QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1174 -GTACGT---GCATTCCTTAATTAATTAAGATTTAACTCGTTCATCAATAATGTGCCAT 1229

QY 187 -----SerLysAlaLysTrpPh 192
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1230 ACCTACTTTAATGTGAATGGTAAGTATGACCCCATCATACAAATAAAGTCATTATGA 1289

QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG 212
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1290 TATGGATCAATATTATATCAAGTTACAAAATTTGATT-----GATAATTTCAGG 1340

QY 212 y-----IleProGluGluPheGlnArgTrpGlnIleAsp 223
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAAATTCCTGAAACTCTACTTAGAT 1392

RESULT 10
US-08-920-827-6
; Sequence 6, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago

```

```

; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
; US-08-920-827-6

Alignment Scores:
Pred. No.: 0.0033 Length: 8654
Score: 114.00 Matches: 62
Percent Similarity: 40.31% Conservatve: 42
Best Local Similarity: 24.03% Mismatches: 97
Query Match: 8.33% Indels: 57
DB: 1 Gaps: 12

US-10-780-002-2 (1-260) x US-08-920-827-6 (1-8654)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTrpSerTrpTrpLys 22
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 681 GAAGGCATCCCATCATTTTAATTCATGGA---TTAGATGGAACCTGGCAGGATTAAA 737

QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 738 GATTATAAATAAGAACTCAAGAGCAG---TATAGAGTAATTAATTAATGATGTCAGAGT 794

QY 42 SerGlyThrAspLeuArgLysIleGluLeuArgThrLeuTrpAspTrpThrLeuPro 61
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 795 CATGGAAATCTTCACGA-----ACAGAATCATATGAATTAAGAATCATGTTGAAGAT 848

QY 62 LeuMetGluLeuMetGluSerLeuAlaAspGluLysValIleLeuValGlyHisSer 81
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 849 TTAATGATTAATGGAGCATTAATATCGATTCGCACATATTTTA---GGACATGAT 905

QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTrpProGlnLysIleTrpAlaAla 101
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 906 ATGGGGGCATCATTCGAGTGAATTTACTGAAATAATCAATATAAAGTGATTACATTG 965

QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 966 ACAATTGTTTCGGCCAAAAGTGAAGACATTCGAAATGGTTTCAACAAATTAATGGTTGAT 1025

QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
   |||||
   : : : : :
   : : : : :
   : : : : :
Db 1026 TACCAAGAAGAATTAGCAGGCTTTAATAATCTGAGGCAATGATTATT----- 1073

```

QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155  
Db 1074 -----TTATCTCTAAATATTTAAAGAGAAAGATAAAGCAATGAATGGGTA 1121  
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLe 175  
Db 1122 TCAAGGCCAAATATATACATAGACCAACTCCGGAAGAAAGTGCAATTGCA----- 1173  
QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186  
Db 1174 -GTACGT---GCATTGCTTAATATTAAGATTAACTCGTGTTCATCATATAATGTGCCAT 1229  
QY 187 -----SerLysAlaLysTyrPh 192  
Db 1230 ACCTACTTTAATGTGAATGGTAAGTATGACCCTCATACAAAATAAAGTCATTATGA 1289  
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG1 212  
Db 1290 TATGATCAATATTATGATCAAGTTACAAAATTGTAATT-----GATAATTCAGG 1340  
QY 212 Y-----IleProGluGluPheGlnArgTrpGlnIleAsp 223  
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAATTCCTGAACTCTACTTAGAT 1392

## RESULT 11

US-08-921-177-6  
; Sequence 6, Application US/08921177  
; Patent No. 5798211  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/921,177  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8654 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus epidermidis  
; STRAIN: Clinical isolate SE-22  
; US-08-921-177-6

Alignment Scores:  
Pred. No.: 0.0033 Length: 8654  
Score: 114.00 Matches: 62  
Percent Similarity: 40.31% Conservative: 42  
Best Local Similarity: 24.03% Mismatches: 97  
Query Match: 8.33% Indels: 57  
DB: 1 Gaps: 12

US-10-780-002-2 (1-260) x US-08-921-177-6 (1-8654)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrpTyrLys 22  
Db 681 GAAGGCATCCCATCATTTTAATTCATGGA---TTAGATGGAACCTGGCAGGATTAAA 737  
QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
Db 738 GATTTAAAAAATGAATCAAGAGCAG---TATAGAGTAATTAATCTTATGATGTCAGAGGT 794  
QY 42 SerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
Db 795 CATGAAAAATCTTCACGA-----ACAGATCATATGAATTAAGATCATGTTGAAGAT 848  
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81  
Db 849 TTAATGATTTAATGGGAGCATTAATATCGATTCGCACATATTTA---GGACATGAT 905  
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101  
Db 906 ATGGGGGCATCATTCGAGTGAATTTACTGAAAAATATCAATATATAAAGTGATACATTG 965  
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121  
Db 966 ACAATTGTTTCGGCCAAAAGTGAAGACATTCGAATGGTTTCAACAAATTAATGTTGAT 1025  
QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141  
Db 1026 TACCAGAGAAATAGCAGGGCTTTAATAAAATCTGAGGCAATGATTATT----- 1073  
QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155  
Db 1074 -----TTATCTCTAAATATTTAAAGAGAAAGATAAAGCAATGAATGGGTA 1121  
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLe 175  
Db 1122 TCAAGGCCAAATATATACATAGACCAACTCCGGAAGAAAGTGCAATTGCA----- 1173  
QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186  
Db 1174 -GTACGT---GCATTGCTTAATATTAAGATTAACTCGTGTTCATCATATAATGTGCCAT 1229  
QY 187 -----SerLysAlaLysTyrPh 192  
Db 1230 ACCTACTTTAATGTGAATGGTAAGTATGACCCTCATACAAAATAAAGTCATTATGA 1289  
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG1 212  
Db 1290 TATGATCAATATTATGATCAAGTTACAAAATTGTAATT-----GATAATTCAGG 1340  
QY 212 Y-----IleProGluGluPheGlnArgTrpGlnIleAsp 223  
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAATTCCTGAACTCTACTTAGAT 1392

## RESULT 12

US-08-362-577C-6  
; Sequence 6, Application US/08362577C  
; Patent No. 5807673  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; US-08-362-577C-6





```
; STRAIN: Clinical Isolate SE-22
US-08-920-828-6

Alignment Scores:
Pred. No.: 0.0033 Length: 8654
Score: 114.00 Matches: 62
Percent Similarity: 40.31% Conservative: 42
Best Local Similarity: 24.03% Mismatches: 97
Query Match: 8.33% Indels: 57
DB: 2 Gaps: 12

US-10-780-002-2 (1-260) x US-08-920-828-6 (1-8654)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTrpSerTrpTyrLys 22
Db 681 GAAGGCATCCCATCATTTTAAATTCATGGA---TTAGATGGAACTTGGCAGGATTTAA 737
QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 738 GATTTAAAGAAATGAATCAAGAACGAG---TATAGAGTAATTTACTTATGATGTCAGAGT 794
QY 42 SerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61
Db 795 CATGGAAATCTTCACGA-----ACAGAATCATATGAATTAAGAGATCATGTTGAAGAT 848
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Db 849 TTAATATGATTTAATGGGAGCAATTAATATCGAATTCGACATATTTTA---GGACATGAT 905
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Db 906 ATGGGGGGCATCATTTGCGAGTGAATTTACTGAAATATATCAATATAAGTGAATTCATGTG 965
QY 102 ValPheLeuAlaPheMetProPheSerValHisAsnSerSerPheValLeuGluGln 121
Db 966 ACAATTGTTTCGCCCAAAAGTGAACACATGCAATGTTTCAACAAATTAATGTTGAT 1025
QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
Db 1026 TACCAAGAAGAATTAGCAGCGCTTTAATAATCTGAGGCAATGATTATT-----1073
QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155
Db 1074 -----TTATCTCTAAATTTTAAAGAAAGATAAGCAATGAATGGGTA 1121
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSeri 175
Db 1122 TCAAGGCCAAAATATATACATAGACCAACTCCGGAAGAAAGTGCAATTGCA-----1173
QY 175 uValArgProSerSerLeuPheMetGluAspLeu-----186
Db 1174 -GTACGT---GCATTGCTTAAATATTAAAGATTAACTCGTGTTCATCATTAATGTGTCCAT 1229
QY 187 -----SerLysAlaLysTyrPh 192
Db 1230 ACTACTTTAATGTGAATGGTAAGTATGACCACTCATCAAAATATAAGTCATATGA 1289
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG 212
Db 1290 TATGGATCAATATATGATCAAGTTACAAAATTTGATTT-----GATAATTCAGG 1340
QY 212 Y-----IleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAATTTCTGAAACTCTACTTAGAT 1392

RESULT 14
US-09-252-991A-9351/c
; Sequence 9351, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9351
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9351

Alignment Scores:
Pred. No.: 7.28-05 Length: 657
Score: 113.00 Matches: 32
Percent Similarity: 46.15% Conservative: 22
Best Local Similarity: 27.35% Mismatches: 45
Query Match: 8.26% Indels: 18
DB: 4 Gaps: 4

US-10-780-002-2 (1-260) x US-09-252-991A-9351 (1-657)

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTrpSerTrpTyrLys 22
Db 610 GACGCAAGCCGCGTATCGCCCTGCATGGCTGGCTGGACAACGCCATGAGTTTCTCGCGC 551
QY 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSer 42
Db 550 CTGGCGCCGAAGCTC-----GCCGGGCTGGTATCGTCCCTCGACTTCGCGGCCAC 497
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArg---ThrLeuTyrAspTyrThrLeuPro 61
Db 496 GGGCATTCGCGCATCGTGGCGGCGGCCAGCTACCTCTCGGACTACGCGCTGGAC 437
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Db 436 GTGCTGATGGTGGCGGAGCAACTGGGCTGG---GAGCGTTTTTCTCTGTTGGGCACTCG 380
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Db 379 ATGGCGGCATC-----GTCTCG 362
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheVal 118
Db 361 GTGCTGCTCGCGGGGCTTCCCGGCGCATCGAGCGGTGGCGGTGATC 311

RESULT 15
US-09-252-991A-9422
; Sequence 9422, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9422
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9422

Alignment Scores:
Pred. No.: 0.00011 Length: 855
Score: 113.00 Matches: 32
```

Percent Similarity: 46.15% Conservative: 22  
 Best Local Similarity: 27.35% Mismatches: 45  
 Query Match: 8.26% Indels: 18  
 DB: 4 Gaps: 4

US-10-780-002-2 (1-260) x US-09-252-991A-9422 (1-855)

```

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
DB 122 GACGGCAAGCCGGTGTATGCCCTGCATGGCTGGACACGCCATGACTTCTCGCGC 181
QY 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAsp 42
DB 182 CTGGCGCGAAGTC-----GCCGGCTGCGTATCGCTCGCTCGACTTCGCCGCCAC 235
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArg--ThrLeuTyrAspTyrThrLeuPro 61
DB 236 GGGCATTCGCGCATCGTCCGAGGCGCCAGCTACTGTCTGGGACTACGCGCTGGAC 295
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
DB 296 GTGCTGATGGTGGCCGAGCAACTGGGCTGG---GACCGTTTTTCCCTGTGGGGCACTCG 352
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
DB 353 ATGGGCGCCATC-----GTCTCG 370
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheVal 118
DB 371 GTGCTGCTCGCGGGGCTTGGCCGAGCGCATCGAGCGGCTGGCGTTGATC 421
    
```

Search completed: September 12, 2005, 22:51:02  
 Job time : 202 secs

**This Page Blank (uspto)**



PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match		24.0%;	Score 218.8;	DB 3;	Length 792;
Best Local Similarity		56.6%;	Pred. NO. 4.1e-60;		
Matches 447;		Conservative 0;	Mismatches 337;	Indels 6;	Gaps 2;
QY	55	GAGACATATGGAGGAGAAACATCATTCTGTTAGTTTCAACGCTTATCATGGAGCCTG	114		
Db	3	GAGTGAGGAGGAGGAGCAACACGTCGTTCTAGTACATGGTCTTGCATGGCGCTG	62		
QY	115	GATCTGGTACAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGGTTACTGTGTGCA	174		
Db	63	GTGCTGGTCAAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGGTAACCCGCGTAGA	122		
QY	175	ACTCGCGCTCCGGGATCGAC---CCAGACCAATCCAGGCGTTGAAACCGTGCAGCA	231		
Db	123	TCTAGCTGCTCCGTTATAGACATGACAGGTCATACAGATATATCCACATGCGMAA	182		
QY	232	ATACTCAAACCGTTGATCGAAACCTCAATCTCTCCAGAGAACGAAAGAGTAACTCT	291		
Db	183	ATACTCAGAGCCATTGATGAGCTAATGACCTCACTTACCAGATGATGAGAAGGTTGTCT	242		
QY	292	GTTTGAATTCAGCTTCGGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGCGAA	351		
Db	243	TGTTGTCATAGCTTAGGAGGTTTGAGTTAGCTATGGCCATGGATATGTTTCCGACCA	302		
QY	352	GATTAAGGTTCTTGTGTTCTCAACGCTTCTTGGCCGACACCAACCCAGTGCTTCTCA	411		
Db	303	AATCTCTGTTCTGTTCTTGTGACTGCTATGATGCGACACCAACCACTCACCATCTT	362		
QY	412	CGTTCTGACAGATATATGGAGATCCCTGGAGGTTTGGAGATTGTGA---GTTTTTCATC	468		
Db	363	CGATGCGGATAGCTTAAGAAAGAAACTTCACGAGAGGAATGTTAGACACCGTGTTTAC	422		
QY	469	TCATGAAACAGAAATGGGACGATGATTTATTTGAAGATGGGACCAAAATTCATGAGGC	528		
Db	423	GAGCGAGAAACCTGATTTCTTCTAGCGAGTTTGGATTTTGGACCAAGATTCATGCCAA	482		
QY	529	ACGCTCTTTACCAAAATTTGTCCTACAGGATTCAGAGCTGGCAAAATGTTGCATAGGCA	588		
Db	483	GAACTTGATCAGTTGTTCTCCAGTCCAGATCTTGAATTTGGCAAAATGTTGTGAGGC	542		
QY	589	AGGGTCATTTTTCACAGAGGATCTATCAAAGAAAGAAAAGTTTTCAGGAGGAGATATGG	648		
Db	543	AAACCCATTGATTAAGAAAGATATGCGAGAGAAAGAGCTTCAGTGAGGAGGATACGG	602		
QY	649	TTCCGTGCAAGGTTTACGTAATGATGATGATGAGACAAAGCCATCCCTCGCATTTCTAT	708		
Db	603	ATCCGTTTACAGCTATATTTATTTGATGCGGAAAGGATCTTGTGTCAACCCGAAAGATTACCA	662		
QY	709	TCGTTGGATGATGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCA	768		
Db	663	GCATCGATGATCAGCACTTTCCCCCAAGAGATTAATGAGAGATCAAGAGCGAGATCA	722		
QY	769	CATGGTGATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGTGTTATTTGCCACGA	828		
Db	723	TATGCCAATGTTCTCCAAGCCTCAACACTATGTGCTCTTCTCTTGGAGATTGCAATAA	782		
QY	829	TTATATGTAA	838		
Db	783	ATATGCTTAA	792		

RESULT 13

ID	ABZ12412
XX	ABZ12412 standard; DNA; 792 BP.
AC	ABZ12412;
XX	
DT	21-JAN-2003 (first entry)
XX	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 217.
XX	
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX	OS	Arabidopsis thaliana.
XX	PN	WO200216655-A2.
XX	XX	
XX	PD	28-FEB-2002.
PF	24-AUG-2001;	2001WO-US026685.
XX	24-AUG-2000;	2000US-0227866P.
PR	26-JAN-2001;	2001US-0264647P.
PR	22-JUN-2001;	2001US-0300111P.
XX	(SCRI )	SCRIPPS RES INST.
PA	(SYGN )	SYNGENTA PARTICIPATIONS AG.
XX	Harper JF,	Kreps J, Wang X, Zhu T;
PI	WPI;	2002-304127/34.
XX		
DR	XX	
XX	PT	Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.
XX	PS	Claim 144; SEQ ID NO 217; 577pp + Sequence Listing; English.
XX	CC	The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office
XX	SQ	Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

Query Match		24.0%;	Score 218.8;	DB 6;	Length 792;
Best Local Similarity		56.6%;	Pred. NO. 4.1e-60;		
Matches 447;		Conservative 0;	Mismatches 337;	Indels 6;	Gaps 2;
QY	55	GAGACATATGGAGGAGAAACATCATTCTGTTAGTTTCAACGCTTATCATGGAGCCTG	114		
Db	3	GAGTGAGGAGGAGGAGCAACACGTCGTTCTAGTACATGGTCTTGCATGGCGCTG	62		
QY	115	GATCTGGTACAGCTCAAGCCCTCTTGAATTCAGCCGCGCCACCGGTTACTGTGTGCA	174		
Db	63	GTGCTGGTACAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGGTAACCGCGTAGA	122		
QY	175	ACTCGCGCTCCGGGATCGAC---CCAGACCAATTCAGGCGGTTGAAACCGTGCAGCA	231		
Db	123	TCTAGCTGCTCCGTTATAGACATGACAGGTCATACAGATATATCCACATGCGAA	182		
QY	232	ATACTCAAACCGTTGATCGAAACCTCAATCTCTCCAGAGAACGAGGATTAATCT	291		
Db	183	ATACTCAGAGCCATTGATGAGCTAATGACCTCACTACAGATGATGAGAAGGTTGTCT	242		
QY	292	GTTTGGATTCAGCTTCGGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGCGAA	351		
Db	243	TGTTGGTATAGCTTAGGAGGTTTGGAGTTAGCTATGGCCATGGATATGTTCCGACCA	302		
QY	352	GATTAAGGTTCTTGTGTTCTCAACGCTTCTTGGCCGACACCAACCCAGTGCTTCTCA	411		
Db	303	AATCTCTGTTCTGTTCTTGTGACTGCTATGATGCGACACCAACCACTCACCATCTT	362		
QY	412	CGTTCTGACAGATATATGGAGATCCCTGGAGGTTTGGAGATTGTGA---GTTTTTCATC	468		
Db	363	CGATGCGGATAGCTTAAGAAAGAAACTTCACGAGAGGAATGTTGTAGACACCGGTTTAC	422		
QY	469	TCATGAAACAGAAATGGGACGATGATTTATTTGAAGATGGGACCAAAATTCATGAGGC	528		

Db 423 GAGCGAGAAACCTGATTTCTTAGCGAGTTTGGATTTTGGACAGAAATTCATGGCCAA 482  
QY 529 ACGTCTTTACCAAAATGTCCTCATAGAGGATTACGAGCTGGCCAAAATGTTGCATAGCA 588  
Db 483 GAACTTGATCAGTTGTTCTCCAGTCCAAGATCTTGAATGGGAAATGTTGGTGAGGCG 542  
QY 589 AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGG 648  
Db 543 AAACCAATGATTAAAGAAAGATATGCGAGAGAGAAAGCTTCACTGAGGAGGATACGG 602  
QY 649 TTCGGTGAACGAGTTTACGTAATCAGTAGTGAACAGAAAGCCATCCCTCGGATTTCA 708  
Db 603 ATCCGTTACAGTATATTTATTTGATGCGGAAAGGATCTTGTGTCACCGGAAGATTACA 662  
QY 709 TCGTTGGGATGATTGATAATTTCAACGCTCTCGAAAGCTACAGAGATCGATGCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCTCCCAAGAAAGTAAATGAGATCAAGAGCGAGATCA 722  
QY 769 CATGGTGATGCTCTCCAAACCCCAAAACTCTTTGCACTCTCTCTCTGTTGCGACCGA 828  
Db 723 TATGCCAATGTTCTCCAGGCTCAACAACTATGTGCTCTCTCTCTGAGATTTGCAATAA 782  
QY 829 TTATATGTAA 838  
Db 783 ATATGCCTAA 792

RESULT 14

ID ADN74620  
XX ADN74620 standard; cDNA; 792 BP.  
AC ADN74620;  
DT 15-JUL-2004 (firet entry)  
XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2515.  
DE gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
KW growth regulator; animal feed product; thale cress;  
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.  
XX Arabidopsis thaliana.  
XX WO2004035798-A2.  
XX 29-APR-2004.  
XX 20-OCT-2003; 2003WO-BP011658.  
XX 18-OCT-2002; 2002EP-00079408.  
XX (CROP-) CROPDISEIGN NV.  
XX Inze D, De Veylder L, Vlieghe K;  
PI WPI; 2004-348466/32.  
XX P-PSDB; ADN74621.

Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.

Claim 1; SEQ ID NO 2515; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up - or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where

CC the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered endoreduplication, biochemistry, signal transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/ or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress cDNA repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa transcription factor, given in an exemplification of the invention.

XX Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

Query Match 24.0%; Score 218.8; DB 12; Length 792;  
Best Local Similarity 56.6%; Pred. No. 4.1e-60;  
Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2;

QY 55 GAGCATATGAGAGGAGAAACATCACTTCGTGTTAGTTTACACACCTTATCATGAGCCTG 114  
Db 3 GAGTGAGAGGAGAGGAGCAACACGTCGTCTTAGTATCATGTCCTTGCCTATGCGCCTG 62  
QY 115 GATCTGTTACAAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGCTGTCGA 174  
Db 63 GTGCTGGTACAAGGTTAAGCGCAGCTCGAGGCTTCTGGCCACCGCTTAACCGCGCTAGA 122  
QY 175 ACTCGCGCCTCCCGGATCGAC---CCACGACCAATCCAGCCCGTTGAAACCGTCGACGA 231  
Db 123 TCTAGTGCCTCCGTTATAGACATGACAGTCAATCACAGATATATCCACATCGGAACA 182  
QY 232 ATACTCAAAACCGTTGATCGAAACCTCAATCTCTCCAGAGAAACGAGGTTAATTC 291  
Db 183 ATACTCAGAGCCATTGATGACGTAATGACCTCACTACAGATGATGAGAAAGTTGTGCT 242  
QY 292 GGTGGATTCAAGTTTCGAGGCGATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAA 351  
Db 243 TGTGGTCATAGCTTAGAGGTTTGTAGTTAGTTAGTATGTCCTATGTTCCGACCAA 302  
QY 352 GATTAAGTCTTCTGTCTTCTCAACGCTTCTTCCCGGACACAAACCGTCCTTCTCA 411  
Db 303 AATCTCTGTTCTGCTCTTTGTGACTGCTATGATGCCAGACACCAAACTCACCATCTTT 362  
QY 412 CGTTCTGGAACAAGTATATGGAGATCGCTGGAGGTTTGGGAGATTGTGA--GTTTTCATC 468  
Db 363 CGTATGGGATAAGCTAAAGAAAGAACTTCCAGAGAGGAATGGTTAGACACCGTGTTCAC 422  
QY 469 TCATGAAACAAGAAATGGGACGATGAGTTTATTCGAAGATGGGACCAAAATTCATGAAGGC 528  
Db 423 GAGCGAGAAACCTGATTTTCTAGCGAGTTTGGATTTTGGACAGAAATTCATGGCCAA 482  
QY 529 ACGTCTTTTACCAAAATTTGCCCATAGAGGATTACGAGCTGGCCAAAATGTTTCATAGCA 588  
Db 483 GAACTTGATCAGTTGTTCTCCAGTCCAAGATCTTGAATTTGGCGAAATGTTGGTGAGGCG 542  
QY 589 AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGG 648  
Db 543 AAACCCATTGATTAAAGAAAGATATGGCAGAGAGAGAAAGCTTCAGTGAGGAGGATACGG 602  
QY 649 TTCGGTCAACGAGTTTACGTAATGAGTAGTGAAGAGCAAAAGCCATCCCTCGCATTTTCAT 708  
Db 603 ATCCGTTACAGTATATTTATTTGATGCGGAAAGATCTTGTCTCACCGGAAGATTACCA 662  
QY 709 TCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGTCTTACAGATCGATGCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTTCCCGCAAGAAAGTAAATGAGATCAAAAGCGCAGATCA 722  
QY 769 CATGGTGATGCTCTCCAAACCCCAAAACTCTTTGCACTCTCTCTCTGCTATGCGACCGA 828  
Db 723 TATGCCAATGTTCTCCAGGCTCAACAACTATGTGCTCTCTCTCTGAGATTTGCAATAA 782  
QY 829 TTATATGTAA 838  
Db 783 ATATGCCTAA 838



Db	783 ATATGCCTAA 792	
Db	RESULT 15	
Db	ABZ13746	
Db	ID ABZ13746 standard; DNA; 792 BP.	
Db	AC	
Db	ABZ13746;	
Db	DT	
Db	21-JAN-2003 (first entry)	
Db	XX	
Db	Arabidopsis thaliana stress regulated gene SEQ ID NO 1551.	
Db	XX	
Db	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
Db	XX	
Db	Arabidopsis thaliana.	
Db	XX	
Db	WO200216655-A2.	
Db	PN	
Db	28-FEB-2002.	
Db	XX	
Db	24-AUG-2001; 2001WO-US026685.	
Db	PF	
Db	24-AUG-2000; 2000US-0227866P.	
Db	PR	
Db	26-JAN-2001; 2001US-0264647P.	
Db	PR	
Db	22-JUN-2001; 2001US-0300111P.	
Db	XX	
Db	(SCRI ) SCRIPPS RES INST.	
Db	PA (SYGN ) SYNGENTA PARTICIPATIONS AG.	
Db	XX	
Db	Harper JF, Kreps J, Wang X, Zhu T;	
Db	XX	
Db	WPI; 2002-304127/34.	
Db	XX	
Db	Identifying a stress condition to which a plant cell has been exposed and	
Db	PT producing plants with increased tolerance to these abiotic stresses.	
Db	XX	
Db	Claim 144; SEQ ID NO 1551; 577bp + Sequence Listing; English.	
Db	XX	
Db	The invention relates to identifying a stress condition to which a plant	
Db	CC cell has been exposed, comprising: (a) contacting nucleic acid	
Db	CC representative of expressed polynucleotides in the plant cell with an	
Db	CC array or probes representative of the plant cell genome; and (b)	
Db	CC detecting a profile of expressed polynucleotides in the plant cell	
Db	CC characteristic of a stress response. The method is useful in the	
Db	CC production of transgenic plants, cells and seeds and in producing plants	
Db	CC with increased tolerance to abiotic stress. The present sequence is that	
Db	CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used	
Db	CC in methods of the invention. Note: The sequence data for this patent is	
Db	CC not represented in the printed specification but is based on sequence	
Db	CC information supplied to Derwent by the European Patent Office	
Db	XX	
Db	Sequence 792 BP; 210 A; 169 C; 187 G; 226 T; 0 U; 0 Other;	
Db	SQ	
Db	Query Match 23.5%; Score 214.2; DB 6; Length 792;	
Db	Best Local Similarity 57.3%; Pred. No. 1.3e-58;	
Db	Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;	
Db	55 GAGACATATGGAGAGAAACATCATTCTGTTAGTTTCACACGCTTATCATGGAGCCTG 114	
Db	3 GAGTGAGGAGAGAGAGAAACATCATTCTGTTAGTTTCACATGTTGCGTGCACGGCGCATG 62	
Db	115 GATCTGTTACAGCTCAAGCCCTCTTGAATCAGCGCCGACCGGTTACTGCTGCTGA 174	
Db	63 GTGCTGGTACAGGTTAAGCCTCTTCTCGAGGCTTTGGGCCCATCGGTAAACCGCCTTGA 122	
Db	175 ACTCGCGGCTCCGGGATCGACCCA---CGACCAATCCAGGCGGTTGAAACCGTCGACGA 231	
Db	123 CCTAGTGTCTTCGGGTATAGACACACACAGGTCATCATCTGATTTCTACATGTGAACA 182	
Db	232 ATACTCAAACCGTTTCAGAAACCCCTCAAAATCTTTCAGAGAACGAGAGGTAATCT 291	
Db	183 ATATTCTGAGCCATTTCAGCAGTAATGACTTCACTTTCATGTCGGAATGATGAGAAAGTTGTA 242	
Qy	292 GGTGGATTTCAGCTTCGGAGGCATCAACATGGCTCTCGCCGCCGACATATATTTCCGCCGAA 351	
Db		
Db	243 CGTTGGTCATAGCTTTGGAGGTTTGAGTTTAGCCCTTAGCCATGGATAAGTTTCCCGATAA 302	
Qy	352 GATTAAGGTTCTTGTTCTCAACGCTTCTTGCCCGACACAAACCCAGCGCTTCTCA 411	
Db		
Db	303 AATCTCTGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362	
Qy	412 CGTTCTGGACAGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTC 465	
Db		
Db	363 CGTCAGGAAAGTTTGCACAGCATGACACACAGAGATGGATGGGCTCTGAGCTCGA 422	
Qy	466 ATCTCATGAAACACAGAAATGGACATGAGTTTATTGAAGATGGACCAAAATTCATGAA 525	
Db		
Db	423 GACATATGTTTCAGATATTCGGGCTTCTCTGTT---CTTCAGCAGGCTTCATGAA 479	
Qy	526 GGCACGCTCTTACCAAAATTTGCCATAGAGGATTAACGAGCTGGCAAAATTTGCGATAG 585	
Db		
Db	480 GCACCGTCTCTACCAACTTTCTCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 539	
Qy	586 GCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATA 645	
Db		
Db	540 GCCTAGTTTCATTTGTTTATTAAATGAAATTCGAAGATGGAGAACTTTCTGAGAAAGGTA 599	
Qy	646 TGGTTCCGGTCAACCGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTT 705	
Db		
Db	600 TGGATCTGTTCTCGAGCTTACATTTGTTGCAAGAGGACACATTAATCTCGAAGACCA 659	
Qy	706 CATTCGTTGGATGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGA 765	
Db		
Db	660 TCAACGATGGATGATCCATAATTATCCGCCGAAATTTAGTGTGATGAGATGGAAGAGACTGA 719	
Qy	766 TCACATGGTGTGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCAC 825	
Db		
Db	720 TCATATGCAATGTTTTCGAAACCTCAACTACTAAGTACCATCTATTTTGGCANTCGCTGA 779	
Qy	826 CGATT 830	
Db	780 CAATT 784	
Search completed: September 12, 2005, 02:21:41		
Job time : 502.745 secs		

**This Page Blank (uspto)**

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 22:54:15 ; Search time 3956.73 Seconds  
 (without alignments)  
 11168.596 Million cell updates/sec

Title: US-10-780-002-36  
 Perfect score: 912  
 Sequence: 1 atcacgctattctcaacaa.....ttttcgggcaacttctc 912

Scoring table: IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.btg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.4	99.8	928	8	AY046015 Arabidops
2	882	96.7	911	8	AY058115 Arabidops
3	808	88.6	808	8	AY142490 Arabidops
4	777	85.2	777	8	AY093714 Arabidops
5	628	68.9	57180	8	ATF18D22 Arabidops
6	341.4	37.4	856	12	AY202287 Arabidops
7	255.2	28.0	891	8	AY751530 Catharant
8	246.8	27.1	1077	8	AF269158 Citrus si
9	241.6	26.5	792	6	AX506625 Sequence
10	241.6	26.5	792	8	BT014881 Arabidops
11	241.4	26.5	935	8	BT002859 Arabidops
12	232.6	25.5	1079	8	AY485932 Nicotiana
13	221.6	24.3	1078	8	HB040402 Hevea bras
14	221.6	24.3	1091	6	AS9586 Sequence 1
15	218.8	24.0	792	6	CQ806104 Sequence
16	218.8	24.0	792	6	AX505522 Sequence
17	218.8	24.0	792	8	BT015031 Arabidops
18	218.8	24.0	878	8	BT014901 Arabidops
19	214.2	23.5	792	6	AX506856 Sequence

20	214.2	23.5	792	8	AY056068 Arabidops
21	214.2	23.5	870	8	AY086590 Arabidops
22	214.2	23.5	932	8	AF361627 Arabidops
23	211	23.1	1021	8	MEAHXN1L M.esculenta
24	210.4	22.1	961	8	AF178576 Rauvolfia
25	205.4	22.5	777	6	CQ829357 Sequence
26	204.8	22.5	1073	8	AK061058 Oryza sat
27	203.4	22.3	928	8	AK117107 Arabidops
28	199.6	21.9	792	8	BT010429 Arabidops
29	199.6	21.9	902	8	AK175772 Arabidops
30	198.6	21.8	777	6	CQ829358 Sequence
31	196	21.5	771	6	BT006227 Arabidops
32	194.4	21.3	771	6	AX506583 Sequence
33	193.4	21.2	770	6	AX651280 Arabidops
34	185.2	20.3	944	8	BT010709 Arabidops
35	183.6	20.1	783	8	BT012414 Arabidops
36	178.6	19.6	795	6	AX654444 Sequence
37	176.8	19.4	924	8	BT004125 Arabidops
38	172.6	18.9	828	8	AY142031 Arabidops
39	172.6	18.9	903	8	AY086981 Arabidops
40	172.6	18.9	904	8	AY074858 Arabidops
41	167.4	18.4	1087	8	BT012867 Lycopersal
42	148.4	16.3	943	8	AK072603 Oryza sat
43	145.2	15.9	948	8	OSPIR7BR Z34270 O.sativa (c
44	140	15.4	123360	8	AC003040 Arabidops
45	139.2	15.3	3948	8	MEAJ3281 Manihot e

## ALIGNMENTS

AY046015	928 bp	mRNA	linear	PLN 18-SEP-2002
Arabidopsis thaliana putative alpha-hydroxynitrile lyase (At5g10300) mRNA, complete cds.				
AY046015				
AY046015.1	GI:15028130			
FLJ CDNA.				
Arabidopsis thaliana (thale cress)				
Arabidopsis thaliana				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.				
1 (bases 1 to 928)				
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.				
Arabidopsis Full Length cDNA Clones				
Unpublished				
2 (bases 1 to 928)				
Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.				
Direct Submission				
Submitted (06-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN Arabidopsis Full-length cDNA): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.				
The Salk, Stanford, PGEC (SSP) Consortium members carried out the				

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

Location/Qualifiers  
 1..928  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="RAFL05-09-N03 (R09923)"  
 /ecotype="Columbia"  
 /note="this clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."  
 gene  
 1..928  
 /gene="At5g10300"  
 5'UTR  
 1..61  
 /gene="At5g10300"  
 62..838  
 /gene="At5g10300"  
 CDS  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative alpha-hydroxynitrile lyase"  
 /protein\_id="AAK76689.1"  
 /db\_xref="GI:15028131"  
 /translation="MERKHFFLVHNAVGHAWIYKLPKLLSAGHRVTAVELAASGI  
 DRPQAVETVDEYSKPLIETLKSIPENEEVILVGSFGGINIALAADIIPAKIKVLV  
 FLNAPLPDTHVPSHLVDKMPGLGDCFSFHHETNGTSLMLKMGPKMKRLQ  
 NCPIEDYSLKMLHQRQFFETEDLSKKEFBEQYGVSVYVMSSEDAIPCPPIRW  
 MIDNFNVKLVIBDGGDHVMVLSKPKQLFDSLALATDYM"  
 misc\_difference 590  
 /gene="At5g10300"  
 /note="compared to genomic sequence resulting in an amino acid sequence difference"  
 /replace="g"  
 839..928  
 /gene="At5g10300"  
 3'UTR  
 /gene="At5g10300"  
 ORIGIN

Query Match 99.8%; Score 910.4; DB 8; Length 928;  
 Best Local Similarity 99.9%; Pred. No. 6.1e-249;  
 Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATCAGCTATTCTCAACAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 60  
 DB 1 ATCAGCTATTCTCAACAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 60  
 QY 61 TATGAGAGGAAACATCACTCGTGTAGTTTACCAACGCTTATCATGAGCGCTGATCTG 120  
 DB 61 TATGAGAGGAAACATCACTCGTGTAGTTTACCAACGCTTATCATGAGCGCTGATCTG 120  
 QY 121 GTACAGGCTCAAGCCGCTCTTGAATACGCGGCCACCGCTTACTGTGTGCAACTCGC 180  
 DB 121 GTACAGGCTCAAGCCGCTCTTGAATACGCGGCCACCGCTTACTGTGTGCAACTCGC 180  
 QY 181 CGCTCCGGGATGACCCAGACCAATCCAGCGCTTGAACCGTCGAGATATCCAA 240  
 DB 181 CGCTCCGGGATGACCCAGACCAATCCAGCGCTTGAACCGTCGAGATATCCAA 240  
 QY 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGAACGAAGAGGTAATCTGGTGGATT 300  
 DB 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGAACGAAGAGGTAATCTGGTGGATT 300

QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTTCGGGGAAGATTAAAGGT 360  
 DB 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTTCGGGGAAGATTAAAGGT 360  
 QY 361 TCTTTGTGTTCTCAACGCTTTCTTCCCGGACACAAACCCACGTCGCTTCTCACGTTCTGGA 420  
 DB 361 TCTTTGTGTTCTCAACGCTTTCTTCCCGGACACAAACCCACGTCGCTTCTCACGTTCTGGA 420  
 QY 421 CAAGTATATGAGATGCGCTGGAGGTTTGGGAGATGTGAGTTTTCATCTCATGAACAAG 480  
 DB 421 CAAGTATATGAGATGCGCTGGAGGTTTGGGAGATGTGAGTTTTCATCTCATGAACAAG 480  
 QY 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 540  
 DB 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 540  
 QY 541 AAATTTGTCCTCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAAAGGTCATTTT 600  
 DB 541 AAATTTGTCCTCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAAAGGTCATTTT 600  
 QY 601 CACAGAGATCTATCAAGAAAGAAAGTTTAGCGAGGAAGATATGTTTCGGTGCACAG 660  
 DB 601 CACAGAGATCTATCAAGAAAGAAAGTTTAGCGAGGAAGATATGTTTCGGTGCACAG 660  
 QY 661 AGTTTACGTAATGAGTAGTGAAGACAAGCCATCCCTCGCATTTTTCATTTCGTTGGATGAT 720  
 DB 661 AGTTTACGTAATGAGTAGTGAAGACAAGCCATCCCTCGCATTTTTCATTTCGTTGGATGAT 720  
 QY 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGTGCT 780  
 DB 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGTGCT 780  
 QY 781 CTCACAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACGATATATATGTAATA 840  
 DB 781 CTCACAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACGATATATATGTAATA 840  
 QY 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGG 900  
 DB 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGG 900  
 QY 901 GCAACTTTCATC 912  
 DB 901 GCAACTTTCATC 912  
 RESULT 2  
 AY058115 Arabidopsis thaliana AT5g10300/F18D22\_70 mRNA linear PLN 04-NOV-2001  
 LOCUS  
 DEFINITION  
 ACCESSION AY058115  
 VERSION AY058115.1 GI:16648678  
 KEYWORDS  
 FLI CDNA.  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE  
 1 (bases 1 to 911)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsis cdna clones  
 Unpublished  
 2 (bases 1 to 911)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## Direct Submission

Submitted (23-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, P0EC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Barth, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

Location/Qualifiers

1..911

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAFL06-14-F12(R13968)"

/ecotype="Columbia"

1..69

70..846

/note="alpha-hydroxynitrile lyase-like protein"

/codon\_start=1

/product="AT5G10300/F18D22\_70"

/protein\_id="AAL25532.1"

/db\_xref="GI:16648679"

/translation="MERKHPFLVHNAYHAGMIWYKLPILLESAGHRVTAVELAAAGI  
DRPQAVTEVSKPLJLQENEEGLVFGGINIALAADIIPAKIKVLV  
FLNAPLPTTHVSHVLDKYMPPGLGDCPESHETRNQFTSLKMGPKMKARLYQ  
NCPEDIELAKMLHQKSFPTDLSEKKEFSEGGYQVRYVMSEDKALPCDFIRW  
MIDNFNVSKYIEIDGDHVMLSKPKQKUFDSLALATDYM"

847..911

3' UTR

## ORIGIN

Query Match 96.7%; Score 882; DB 8; Length 911;  
Best Local Similarity 99.3%; Pred. No. 8.2e-241;  
Matches 896; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ATCACGCTATTCTCAACCAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACA 60  
DB |||  
DB 10 ATTACGCTATTCTCAACCAAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGAC- 68  
QY 61 TATGGAGAGAAACATCATCTCGTGTAGTTTCAACACGTTTATCATGGAGCTTGGATCTG 120  
DB |||  
DB 69 TATGGAGAGAAACATCATCTCGTGTAGTTTCAACACGTTTATCATGGAGCTTGGATCTG 128  
QY 121 GTACAGCTCAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGCTGCTGAACTCGC 180  
DB |||  
DB 129 GTACAGCTCAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGCTGCTGAACTCGC 188  
QY 181 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGGTTGAAACCGCTGACGAATATCCCAA 240  
DB |||  
DB 189 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGGTTGAAACCGCTGACGAATATCCCA 248  
QY 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGACGAAAGGTAATTCGTTGGATT 300  
DB |||  
DB 249 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGACGAAAGGTAATTCGTTGGATT 308

QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTTCGGCGCAAGATTAAGGT 360  
DB |||  
DB 309 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTTCGGCGCAAGATTAAGGT 368  
QY 361 TCTTGTGTTCTCAACGCTTCTTCCCGACACAAACCCACGCGCTTCTCAGCTTCTGA 420  
DB |||  
DB 369 TCTTGTGTTCTCAACGCTTCTTCCCGACACAAACCCACGCGCTTCTCAGCTTCTGA 428  
QY 421 CAAATATATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACACAG 480  
DB |||  
DB 429 CAAATATATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACACAG 488  
QY 481 AAATCGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCAACGCTTTTACCA 540  
DB |||  
DB 489 AAATCGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCAACGCTTTTACCA 548  
QY 541 AAATGTTCCCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTT 600  
DB |||  
DB 549 AAATGTTCCCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTT 608  
QY 601 CACAGAGATCTATCAAGAGAAAGTATTTAGCGAGGAGGATATGTTCCGTCGCAAGC 660  
DB |||  
DB 609 CACAGAGATCTATCAAGAGAAAGTATTTAGCGAGGAGGATATGTTCCGTCGCAAGC 668  
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT 720  
DB |||  
DB 669 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT 728  
QY 721 TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATCGCGAGATCACATCGTGATGCT 780  
DB |||  
DB 729 TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATCGCGAGATCACATCGTGATGCT 788  
QY 781 CTCGAAACCCCAAAACCTTTTGACTCTCTCTGCTATTCGACCGGATTAATGTAATA 840  
DB |||  
DB 789 CTCGAAACCCCAAAACCTTTTGACTCTCTCTGCTATTCGACCGGATTAATGTAATA 848  
QY 841 ATCTTAAGTTCGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900  
DB |||  
DB 849 ATCTTAAGTTCGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 908  
QY 901 GC 902  
DB |||  
DB 909 GC 910

AY142490 808 bp mRNA linear PLN 23-SEP-2002  
Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
(AT5G10300) mRNA, complete cds.  
ACCESSION AY142490  
VERSION AY142490.1 GI:23296321  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1 (bases 1 to 808)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T.,  
Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G.,  
Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 808)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T.,  
Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G.,  
Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE Direct Submission  
JOURNAL Submitted (16-AUG-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
COMMENT Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.  
FEATURES  
source  
1..808  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="C104719"  
/ecotype="Columbia"  
/note="this clone is in pUNI 51."  
1..808  
/gene="At5g10300"  
1..777  
/gene="At5g10300"  
/codon\_start=1  
/evidence=experimental  
/product="putative alpha-hydroxynitrile lyase"  
/protein\_id="AA113041.1"  
/db\_xref="GI:23296322"  
/translation="MERKHFVLVHNAVHGAWIYKLPLESAGHVRTVAELAAGSI  
DPRPQAVETVDEYSKPLIETLSLPENEVILVGFSGGINIALAADIPFAKIKVLV  
FLNAFLPDTTHVPVSLDKMEMPGGLGDCFSFSEHTRNGMSLLKMGPKFMKARLVQ  
NCPIEDYELAKMLHRQGSFFTEDLSKKEFSEEGYSVORVYVMSSEDKAIPCDFIRW  
MIDNFNVSKVYIEIDGDHVMVLSKPKLFDLSLSAATDYM"  
778..808  
/gene="At5g10300"  
3' UTR  
ORIGIN  
Query Match 88.6%; Score 808; DB 8; Length 808;  
Best Local Similarity 100.0%; Pred. No. 1.2e-219;  
Matches 808; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 62 ATGGAGGGAACATCACTTCGTGTAGTTCACACGCTTATCATGGAGCGCTGGATCGG 121  
DB 1 ATGGAGGGAACATCACTTCGTGTAGTTCACACGCTTATCATGGAGCGCTGGATCGG 60  
QY 122 TACAAGCTCAAGCCCTCTTGAATCAGCGGCCACCGCGTGTACTGTGTGCAACTCGCC 181  
DB 61 TACAAGCTCAAGCCCTCTTGAATCAGCGGCCACCGCGTGTACTGTGTGCAACTCGCC 120  
QY 182 GCCTCGGATCGACCCAGCAATCCAGCGCGTGTGAAACCGTCGACGAATACCTCCAAA 241  
DB 121 GCCTCGGATCGACCCAGCAATCCAGCGCGTGTGAAACCGTCGACGAATACCTCCAAA 180  
QY 242 CCGTTGATCGAACCTCCTCAATCTCTCCAGAGACGAGAGTAACTCTGGTTGGATTC 301  
DB 181 CCGTTGATCGAACCTCCTCAATCTCTCCAGAGACGAGAGTAACTCTGGTTGGATTC 240  
QY 302 AGCTTCGAGGATCAACATCCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTT 361  
DB 241 AGCTTCGAGGATCAACATCCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTT 300  
QY 362 CTTGTGTTCTCAAGCCCTTTTGGCCGACACAAACCGCGTGTCTCAAGTTCTGGAC 421  
DB 301 CTTGTGTTCTCAAGCCCTTTTGGCCGACACAAACCGCGTGTCTCAAGTTCTGGAC 360  
QY 422 AGGTATGAGATCGCTGGAGTTTGGAGATCTGAGTTTTCATCTCATGAAACAAGA 481  
DB 361 AGGTATGAGATCGCTGGAGTTTGGAGATCTGAGTTTTCATCTCATGAAACAAGA 420  
QY 482 AATGGACGATGAGTTTATGAGATGGGACCAAAATTCATGAGGCGAGCTTTTACCAG 541  
DB 421 AATGGACGATGAGTTTATGAGATGGGACCAAAATTCATGAGGCGAGCTTTTACCAG 480  
QY 542 AATGTGCCATGAGATTACAGCTGGCAAAATATGTTGCATAGGCAAGGTCATTTTC 601  
DB 481 AATGTGCCATGAGATTACAGCTGGCAAAATATGTTGCATAGGCAAGGTCATTTTC 540  
QY 602 ACAGAGGATCTATCAAGAAAGAAAGTTTACGGAGGAGGATATGGTTTCGTCGCAACGA 661

Db 541 ACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGGTTTCGTCGCAACGA 600  
QY 662 GTTTACGTAATGACTAGTGAAGACAAAGCCATCCCTGCGGATTTTCATTCGTTGATGATT 721  
Db 601 GTTTACGTAATGACTAGTGAAGACAAAGCCATCCCTGCGGATTTTCATTCGTTGATGATT 660  
QY 722 GATAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCATCATGGTATGCTC 781  
Db 661 GATAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCATCATGGTATGCTC 720  
QY 782 TCCAAACCCCAAAACTCTTTGACTCTCTCTCTATTGCCACCGATTATATGTAATAA 841  
Db 721 TCCAAACCCCAAAACTCTTTGACTCTCTCTCTATTGCCACCGATTATATGTAATAA 780  
QY 842 TCTTAAGTCCGTTTACTTTTTTCTCAT 869  
Db 781 TCTTAAGTCCGTTTACTTTTTTCTCAT 808  
AY093714 777 bp mRNA linear PLN 13-APR-2002  
Arabidopsis thaliana AT5g10300/F18D22\_70 mRNA, complete cds.  
DEFINITION  
ACCESSION AY093714  
VERSION AY093714.1 GI:20147248  
KEYWORDS  
SOURCE FLI CDNA.  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 777)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,  
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,  
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,  
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,  
Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,  
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,  
Theologis, A. and Ecker, J.R.  
Arabidopsis ORF clones  
Unpublished  
2 (bases 1 to 777)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,  
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,  
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,  
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,  
Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,  
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,  
Theologis, A. and Ecker, J.R.  
Direct Submission  
Submitted (29-MAR-2002) Salk Institute Genomic Analysis Laboratory  
(SIGnAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.  
The Salk, Stanford, PGENC (SSP) Consortium members constructed and  
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Shinn, P.,  
Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,  
Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,  
Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S.,  
Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M.,  
Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W.,  
Theologis, A., and Ecker, J.R.  
Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (ssp/salk) contributed equally to this work as PIs.

FEATURES	source	Location/Qualifiers
		1. .777
		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
		/chromosome="5"
		/clone="U13968"
		/ecotype="Columbia"
CDS		1. .777
		/note="alpha-hydroxynitrile lyase-like protein"
		/codon_start=1
		/product="AT5g10300/P18D22_70"
		/protein_id="AAU10338.1"
		/db_xref="GI:20147249"
		/translation="MERKHPVLVHNAYHGNWIYKLPILLESAGHRVTAVELAASGI DPRIPAVETVDEYSKLIETKSLPENEEVILVGFSGGINIALAADIPPAKIKVLV FLNAPLDTHTVPSHVLDKMEMPGGLGDECFSSHSTRNGTMSLLKMGPKPMKARLYQ NCTPEDYELAKMLHROGSPTFEDLSKKEFSSEEGYGVQRYVYVMSSEDKAIPCDFIRW MIDNFNVSKVVEIDGGDHVMWLSKPKLFDLSAIAIDYM"
ORIGIN		
	Query Match	85.2%; Score 777; DB 8; Length 777;
	Best Local Similarity	100.0%; Pred. No. 8.9e-211;
	Matches 777; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	62	ATGCAGAGGAAACATCACTCGTGTAGTTTCACACCGCTTATCATGAGCGCTCGATCTGG 121
Db	1	ATGCAGAGGAAACATCACTCGTGTAGTTTCACACCGCTTATCATGAGCGCTCGATCTGG 60
Qy	122	TACAAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGCTGCAACTCGCC 181
Db	61	TACAAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGCTGCAACTCGCC 120
Qy	182	GCCTCCGGGATCGACCCAGCACCAATCCAGGCGCTTGAAACCGTCGAGCAATATCCCAA 241
Db	121	GCCTCCGGGATCGACCCAGCACCAATCCAGGCGCTTGAAACCGTCGAGCAATATCCCAA 180
Qy	242	CCGTTGATCGAAACCCCTCAAAATCTCTTCACAGAACGAAGAGGTAATCTGTTGGATTC 301
Db	181	CCGTTGATCGAAACCCCTCAAAATCTCTTCACAGAACGAAGAGGTAATCTGTTGGATTC 240
Qy	302	AGCTTCGAGGCATCAACATCGCTCTGCGCGCGACATATTTCCGCGAAGATTAAAGTT 361
Db	241	AGCTTCGAGGCATCAACATCGCTCTGCGCGCGACATATTTCCGCGAAGATTAAAGTT 300
Qy	362	CTTGTGTTCTCAACGCCTCTTCCCGACACCAACCGTGCCTTCTCAGTTCTGAC 421
Db	301	CTTGTGTTCTCAACGCCTCTTCCCGACACCAACCGTGCCTTCTCAGTTCTGAC 360
Qy	422	AAGTATATGGAGATGCCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAGA 481
Db	361	AAGTATATGGAGATGCCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAGA 420
Qy	482	AATCGGACGATGAGTTTATGAAGTGGACCAAAATTCATGAGGCACGCTCTTACCA 541
Db	421	AATCGGACGATGAGTTTATGAAGTGGACCAAAATTCATGAGGCACGCTCTTACCA 480
Qy	542	AATTGTCCCATAGAGGATTACGAGCTGGCAAAAATGTTGTCATAGGCAAGGTCATTTTTC 601
Db	481	AATTGTCCCATAGAGGATTACGAGCTGGCAAAAATGTTGTCATAGGCAAGGTCATTTTTC 540
Qy	602	ACAGAGATCTATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTTCGGTCAACGA 661
Db	541	ACAGAGATCTATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTTCGGTCAACGA 600
Qy	662	GTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGGATGATT 721
Db	601	GTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGGATGATT 660
Qy	722	GATAAATTCAACGCTCTCGAAAGTCTACGAGATCGATCGCGAGATCACTGGTGATGCTC 781









ISVAFVFTAMPDTHRPVLEQYSEKMKEDDSDLWTFQSCDASNPISHI.SMLFGR  
EFLTIKIYOLCPEDLELAKMLVRPGSMFIDNLKSKFSDGYSKRVLYLVCBEDI  
GLPKQFQHWMIQNPVNEWIEIKGGDHMAWLSDPQKLCCLSLISUKYA"

## ORIGIN

Query Match 27.1%; Score 246.8; DB 8; Length 1077;  
Best Local Similarity 58.5%; Pred. No. 5e-59;  
Matches 456; Conservative 0; Mismatches 312; Indels 12; Gaps 1;  
ORIGIN

QY 75 ATCACTTCGTTAGTTCAACGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGC 134  
DB 83 AGCATTTTGTCTAGTTTCATGGAGTAACCATGAGCATGGTGTGGTACAACTGAAAG 142  
QY 135 CCCTCTCTGAATCAGCGCGCACCGGCTTACTCTGTGCAACTCGCGGCTCTCGGATCG 194  
DB 143 CAAGGCTGGTGGCGGGGGTCAACGGGTGACGCTGTGGACCTAGCGGCTCGGGATCA 202  
QY 195 ACCCAGCAATTCACAGCGGTTGAAACCGTGCAGCAATCTCCTCAAAACGTTGATCGAAA 254  
DB 203 ACATGAAGAGAAATTGAGGATGTGCACACATTCATGCATACAGTGAAGCCCTTGATGGAG 262  
QY 255 CCCTCAAACTCTTCAGAGAACGAGAGTAATCTGGTGGATTCGATTGAGTTCAGCTTCGGAGGCA 314  
DB 263 TTTTGGGATCAGTTCCTGCGGAGAAAGGTCTACTTGTGCGACACAGTCTTGGGGGG 322  
QY 315 TCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTGTCTTGTGTTCTCTCA 374  
DB 323 TCATTTGGCCCTTGGCGGACAAATTCACACAAATCTCCGTGGCTGTTCGTAA 382  
QY 375 AGCCCTTTTGGCCGACAAACCCACGTCCTTCTCACGTTCTGGCAAGATATATGGAGA 434  
DB 383 CTGCATTTCATGCTGCACACACACCGGCCATCTTTGTTTGGAGCAGTATCTCTGAGA 442  
QY 435 TGCCTGGA-----GGTTTGGAGATTGTGAGTTTTCATCTCATGAAACGAGAA 482  
DB 443 AGATGGGAAAGAGGACGACAGCTGTTGGACACTCAATTTTTCACAAATGTGACGCGTCAA 502  
QY 483 ATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTTACCAA 542  
DB 503 ATCCATCTCACATTTCCATGCTTTTCGGCGGAGTCTTGACTATCAAGATCTATCAGC 562  
QY 543 ATTGTCCCATAGAGGATACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTCA 602  
DB 563 TTTGTCTCTGAGGATCTGAGCTGGCCCAAGATGTTGGTGGCCAGGATCAATGTTTA 622  
QY 603 CAGAGGATCTATCAAGAAGAAAGTTTACGAGGAAGGATATGTTTCGGTGCACACGAG 662  
DB 623 TAGACAACTTATCGAAGGAAAGTAAGTTTCAGCGATGAAGGATACCGGATCTGTTAAGCGAG 682  
QY 663 TTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGATG 722  
DB 683 TTTTATCTTGTATGCGAAGAGGATATTTGGTCTCCCTAAGCAATTCAGCACTGGATGATCC 742  
QY 723 ATAATTTCAAGCTCTCGAAGTCTACAGATCGATGGCGGAGATCAATGGTGTGCTCT 782  
DB 743 AAAACTATCCGTTAATGAAGTATGATGAGATCAAGGGCGGTGATCACATGGCAATGCTTT 802  
QY 783 CCMAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACCGATTTATATGTAATAT 842  
DB 803 CCGACCCACAGAAACTTTCCGATTTGTCTGTCTCAGATTTCTTCTTAAGTATGCATAAAAT 862

## RESULT 9

AX506625  
LOCUS AX506625 792 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 1320 from Patent WO0216655.  
ACCESSION AX506625  
VERSION AX506625.1 GI:23387862  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.  
1  
REFERENCE  
AUTHORS Harper, J. F., Kreps, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 1320 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG (CH)  
FEATURES  
source 1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 26.5%; Score 241.6; DB 6; Length 792;  
Best Local Similarity 59.6%; Pred. No. 1.4e-57;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCACCTCGTGTAGTTTCAACGCTTATCATGGAGCCTGGATCTGGTACA 125  
DB 14 AGAGGAAACATTTTGTACTAGTACATGTTCTGTCATGGCGGTGGTCTGGTACA 73  
QY 126 AGCTCAAGCCCTCCTTGAATCAGCCGCGCACCGGTTACTGCTGTGAACTCGCGCT 185  
DB 74 AGGTTAAGCCGCTGTAGAGCGGTGGGCGCACCGGTAACCTGCTGGACTTAGCTGCCT 133  
QY 186 CCGGATCGACCCA---CGACCAATCCAGGCGGTGAAACCGTCGACGATCTCAAC 242  
DB 134 CCGGAATAGACAAACGAGGTGATCTGACATCCCCACATCGCAACATATCTGGAGC 193  
QY 243 CGTTGATCGAAACCCCTCAAACTCTTCCAGAGAACGAGAGTAATCTTGGTTGGATTCA 302  
DB 194 CATTGACGAGCTCTTGACCTCATTTGCCAATGATGAAAGGTTGCTGCTGGTACA 253  
QY 303 GCTTCGAGGACATCAACATGCTCTCGCGCGGACATATTTCCGCGCAAGATTAAAGTTTC 362  
DB 254 GCTTTGGTGGCTTGAACCTTAGCCATAGCATGGAAGTTTCCGGAAGAAATCTCTGTG 313  
QY 363 TTGTGTTCTCAACGCTTCTTGGCCGACACACCCAGTCGCTTCTCACGTTCTGGACA 422  
DB 314 CTGTATTTCTGACTGCTTTTCATGCGGACACCGAACATCTCACCATCTTCTGCTTGGACA 373  
QY 423 AGTAT-----ATGGAGATGCTGAGGAGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA 476  
DB 374 AGTTTGGAGACAACTGCTCAAGAGCATGGATGGGACCCGAAATTCGAACTTATGGTT 433  
QY 477 CAAGAAATGGGACGATGAGTTTATTTGAAGTGGGACCAAAATTCATGAAGGACGCTTTT 536  
DB 434 CAGACAAATTCGGACTGAGTATGTT---TTTAGCCCTGACTTTCATGAAGTTGGGTCTCT 490  
QY 537 ACCAAATTTGCCATAGAGATTACGAGCTGGGCAAAATGTTGATAGGCAAGGTCAT 596  
DB 491 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATAGGCCAGGATCGT 550  
QY 597 TTTTTCACAGAGATCTATCAAGAAAGAAAGTTTTCGAGGAGGAGATATGTTTCGGTGC 656  
DB 551 TATTTATTAACGATTTATCGAAGATGAAAACTTCTCGATGAAGGATATGGGTCTGTTTC 610  
QY 657 AACGAGTTTACGTAATGAGTAGTGAAGCAAGCCATCCCTTGGCATTTCAATTCGTTGGA 716  
DB 611 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAACCGCCAGAGATGGA 670  
QY 717 TGATGTAATTTCAACGCTCTGAAAGTCTAGGATCGATGGCGGAGATCAATCATGGTGA 776  
DB 671 TGATTTGTAATTTTTCGGGTGAATTTTAGTGTGGAGATGGAGAGACAGATCATATGCCAA 730  
QY 777 TGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGCATATATGT 836  
DB 731 TGTCTGCAAGCTCAGCACTCAGTATTACTTCTGAAATTCGCGACAAATTCGTTT 790  
QY 837 AA 838  
||

Db	791 AA 792	
RESULT 10		
BT014881	792 bp mRNA linear	PLN 05-JUN-2004
LOCUS	Arabidopsis thaliana At2g23620 gene, complete cds.	
DEFINITION	Arabidopsis thaliana	
ACCESSION	BT014881	
VERSION	BT014881.1	GI:48310670
KEYWORDS	FLI CDNA.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	Chen, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.	
TITLE	Arabidopsis ORF clones	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 792)	
AUTHORS	Chen, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-JUN-2004) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
FEATURES	Location/Qualifiers	
source	1..792	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/db_xref="taxon:3702"	
	/chromosome="2"	
	/clone="U67545"	
	/ecotype="Columbia"	
	/note="This clone is in puni 51"	
CDS	1..792	
	/note="unknown protein"	
	/codon_start=1	
	/product="At2g23620"	
	/protein_id="AA141864.1"	
	/db_xref="GI:48310671"	
	/translation="MSERKKQHFVLVHSGCMVCKVPLLEAVGHRTAVDLAA SGIDTRITDPTCEQSEPLTKLLSLPNDEKVLVGHSPFGLNLALAMEKPEKI SVAVFLTAPMPTEHSPFLDKFSNMPQEAAMGTEFEPEYSDNSGLSMFSPDFMK LGLYQLSPVEDLELGLLMRPGSLFINDLSKKNFSDSEYGVSPRVFVCKEDKAIFE ERQWMIDNFPVNLVMEBETDHPMFCKPQOLSDFLKIADKFV"	
ORIGIN		
Query Match	26.5%; Score 241.6; DB 8; Length 792;	
Best Local Similarity	59.6%; Pred. No. 1.4e-57;	
Matches	466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;	
Qy	66	AGAGGAAACATCACTTCGGTGTAGTTTCAACAGCTTATCATCGAGCTGGATCTGGTACA 125
Db	14	AGAGGAAACACATTTGTACTAGTACATGTTGTGTCATCGCGGTGCTCTGTGTTACA 73
Qy.	126	AGCTCAAGCCCTCTTGAATCAGCGGCCACCGGTTACTCTGTGCAACTCGCGGCT 185
Db	74	AGGTTAAGCCGCTGTAGAGCGGTGGGCCACCGGTAAGTCTGTGGACTTAGCTGCT 133
Qy	186	CCGGATCAACCA---CGACCAATCCAGCGCTTGAACCCGTCGACGATCTCAAC 242
Db	134	CCGGAATAGACACACGAGGTGATCATCTGATCCCATCCCCATCGCAACATACTCGGAGC 193
Qy	243	CGTTGATCGAAACCCCTCAAACTCTTCCAGAGAACGAGAGTAACTCTGGTTGGATTCA 302
Db	194	CATTGACGAGCTCTGACTCATTCGCAAAATGATGAAGGTTGTCTCTGTGTTACA 253
Qy	303	GCTTCGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAGGTTTC 362
Db	254	GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGMAAAGTTTCCCGAAAAAATCTCTGTGCG 313
Qy	363	TTGTGTCTCAAGCGCTTCTTCCCGCAGACACACCAACCGTGCCTTCTCAGGTTCTGGACA 422
Db	314	CTGTATTCTTCACTGCTTTCAATGCCGACACCGCAACTCAACATCTCTGCTTCTGGACA 373
Qy	423	AGTAT-----ATGGAGATGCTGGAGGTTTGGGAGATTTGTAGTTTTTCATCTCATGAAA 476
Db	374	AGTTTGGAGCAACATGCCTCAAGAAGCATGGATGGGCACCGAATTCGAACCTTTATGGTT 433
Qy	477	CAAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAGGCACGTCCTTT 536
Db	434	CAGACAAATTCGGACTCGATGATGTT---TTTATGCCCTGACTTCATGAAGTTGGGTCTCT 490
Qy	537	ACCAAAATTTGCCATAGAGGATTACGAGCTGGGCAAAATCTTGCATAGGCAAGGTCAT 596
Db	491	ACCGAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTAATGAGGCCAGATCGT 550
Qy	597	TTTTACAGAGGATCTATCAAAAGAAAGAAAGTTTACGAGGAAAGGATATGTTTCGGTGC 656
Db	551	TATTATTAACGATTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGTCTGTTTC 610
Qy	657	AACGAGTTTACGTAATGAGTAGTGAAGCAAGCATCCCTCGGATTTCAATCGTTGGA 716
Db	611	CTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATTCAGAGAAAGCCGAGAGATGA 670
Qy	717	TGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGTTGA 776
Db	671	TGATTGATAATTTTCCCGGTGAATTTAGTAGTGGAGATGGAGAGACAGATCATATGCCAA 730
Qy	777	TGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGATTCGCCACCGATTATATGT 836
Db	731	TGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTTGAAAATCGCGACAAATTCGTTT 790
Qy	837	AA 838
Db	791	AA 792
RESULT 11		
BT002859	935 bp mRNA linear	PLN 15-JAN-2003
LOCUS	Arabidopsis thaliana clone RAPL14-80-F10 (R20356) putative acetone-cyanohydrin lyase (At2g23620) mRNA, complete cds.	
DEFINITION	Arabidopsis thaliana	
ACCESSION	BT002859.1	GI:27754456
VERSION	BT002859.1	
KEYWORDS	FLI CDNA.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.	
TITLE	Arabidopsis Full Length cDNA Clones	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 935)	
AUTHORS	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL CDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,	

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome

#### FEATURES

Location/Qualifiers

1..935

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="2"

/clone="RAFL14-80-F10 (R20356)"

/ecotype="Columbia"

/notes="This clone is in a modified pBluescript vector2

(lambda PS) as a BamHI/XhoI insert."

1..935

/gene="At2g23620"

1..43

/gene="At2g23620"

44..835

/gene="At2g23620"

/codon\_start=1

/evidence=experimental

/product="putative acetone-cyanohydrin lyase"

/protein\_id="AAO22676.1"

/db\_xref="GI:27754457"

/translation="MSEERKKOHLVHSGHGWKVKPLLEAVGHRVTAVDLAA

SGIDTFRSLDIPTEQYSEPLKLLSLPNDEKVLVGHSEGLNLALAMEKFPKLI

SVAVFLATMPDTEHSFSLDKFGSNMPEQAWMGTEFEPYGDNSGLSMFSPDFMK

LGLYQLSPDELELGLMLRPGSLFINDLSKMNFSDEGYSVPRFVIVCKEDKAIFE

ERQRWIDNFPVNLVMEEMETDHPMFCKPQQLSDYFLKADKPV"

misc\_difference 341

/gene="At2g23620"

/note="compared to genomic sequence resulting in an amino

acid sequence difference"

/replaces="g"

836..935

/gene="At2g23620"

ORIGIN

Query Match 26.5%; Score 241.4; DB 8; Length 935;

Best Local Similarity 59.5%; Pred. No. 1.7e-57;

Matches 467; Conservative 0; Mismatches 306; Indels 12; Gaps 3;

QY 66 AGAGAAACATCACTTCGTGTAGTTTCACACGCTTATCATGAGCGTGGTACA 125

DB 57 AGAGAAACAACTTTGTACTAGTACATGGTTTCGTCATGCGCGTGGTACA 116

QY 126 AGCTCAAGCCCTTCCTTGAATCAGCGCGCACCGCGTTACTGTCGCAACTCGCGCCT 185

DB 117 AGTTTANGCGGTGTAGAGCGGTGGCCACCGGTAACTGCTGTGGATTAGTGTCT 176

QY 186 CCGGGATCGACCCCA---CGACCAATCCAGCGGTGAAACCGTCGACGAATACTCAAC 242

DB 177 CCGGAATAGACACACGAGTCGATCACTGACATCCCATCCCATGCGCAATACTCGGAC 236

QY 243 CGTTGATCGAAACCTTCAATCTCTCCAGAGAACGACAGTAATCTGGTTGATTCA 302

DB 237 CATTTGCAAGTCTCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTGTGTCACA 296

QY 303 GCTTCGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTC 362

Db 297 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAAGTTTCCCAAAAAATCTCTGTGC 356

QY 363 TTGCTCTTCTCAACCCCTTCTTGGCCGACACACCCAGTGGCTTCTCACGTTCTGGACA 422

Db 357 CTGTATTTCTTGACTCTCTTTCATGCGCGACACCCGAACACTTCACATCTTCTGGTCTGGACA 416

QY 423 AGTAT-----ATGAGAGATGCTCGAGAGTTTGGAGATTTGTGAGTTTTTTCATCTCATGAAA 476

Db 417 AGTTTGGAGCAACATGCTCAAGAAGCATGATGGCACCGCAATTCGAACCTTATGGTT 476

QY 477 CAAGAAATGGGACGATGATTTATTTGAAGATGGGACCAAAATTCATGAAGGACGCTTTT 536

Db 477 CAGACAAATTCGCGACTGAGTATGTT---TTTATGCGCTGACTTCATGAAGTTGGGTCTCT 533

QY 537 ACCAAATTTGCTCCCATAGAGGATACGAGCTGCGCAAAATTTTGCATAGGCAAGGTCAT 596

Db 534 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGATTAATTTAATAGGCCGAGGATCT 593

QY 597 TTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTATAGCGAGGAGGATATGTTTCGGTGC 656

Db 594 TATTTATTAAAGATTTATCGAAGATGAAAGAACTTCTCGATGAAGATATGCGTCTGTTTC 653

QY 657 AACGAGTTTACATAATGATAGTAGTGAAGACAAAGCCATCCCTCGCGATTTTCATTCGTTGA 716

Db 654 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAACGCCAGAGATGGA 713

QY 717 TGATTTGATATTTTCAACGCTCTCGAAAGTCTAGAGATCGATGGCGAGATCATACATGGA 776

Db 714 TGATTTGATATTTTCCCGTGAATTTAGTAGTGAGATGGAGAGACAGATCATATATATGCT 773

QY 777 TGCTCTCCAAACCCCAAAACCTCTTTGACTCTCTCTGCTATTTGCCACCGGATATATGCT 836

Db 774 TGTCTCGAAGCTCAGCACTCAGTGATTTACTTCTGAAATATCGGACAAATTCGTTT 833

QY 837 AATAA 841

Db 834 AATCA 838

AY485932 1079 bp mRNA linear PLN 07-JAN-2004

Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.

AY485932

AY485932.1 GI:40549302

Nicotiana tabacum (common tobacco)

Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana.

Kumar,D. and Klessig,D.F.

1 (bases 1 to 1079)

High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity

Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)

14673096

2 (bases 1 to 1079)

Kumar,D. and Klessig,D.F.

Direct Submission

Submitted (24-NOV-2003) BFI, Tower Rd., Ithaca, NY 14853, USA

Location/Qualifiers

1..1079

/organism="Nicotiana tabacum"

/mol\_type="mRNA"

/cultivar="Xanthi nc [NN]"

/db\_xref="taxon:4097"

42..824

/note="lipase; SAPP2"

/codon\_start=1

CDS

/product="ealicylic acid-binding protein 2"  
/protein\_id="AAR87711.1"  
/db\_xref="GI:40549303"  
/translation="MKEGKHFVLVHGACGGMSWYKLPLEAAGHKVATDLAASGT  
DLRKLELRTLYDTLPWELMESLADENKVLTVGSLGSMNLGLAMEKPKFIYAAY  
FLAAMPDSVHNSFVLBOYNERTPAENKLDTPLPYGPSPEPLTSMFPGPKFLAHL  
YQLCSPEDLALASSIVRPSLSFPMEDLSKAKYFTDPRFGSVKRVYIVCTEDKGIPEFO  
RMQIDNVGTEAIEIKGADHMMUCEPKQLCASLLEIAHKYN"

ORIGIN	Query Match			
	Best Local Similarity	Score	DB 8;	Length 1079;
	Matches 458;	Conservative 0;	Mismatches 349;	Indels 6; Gaps 1;
QY	33	AACCAAAAGAAAAGCTTACCGGACATATCGAGAGGAAACATCACCTGGTGGTTAGTTC	92	
Db	13	AAAGAAACATAAAGGCGTAAATTCAAATGAGGAAGGAAACACATTTGTTTGTAGTAC	72	
QY	93	ACAAAGCTTATCATGGAGCTGGATCTGGTACAAGCTCAAGCCCTCTTGAATCAGCG	152	
Db	73	ATGGTGATGCCATGGAGGTGGATGGTACAAGCTAAAGCCACTGCTAGAAAGCTGAG	132	
QY	153	GCACCGCGTTACTGCTGCGAACTCGCGCTCCGGGATCGACCAAGCAATCCAGG	212	
Db	133	GCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGTATTTGAGAAAATAGAGG	192	
QY	213	CCGTTGAAACCGTCGACGCAATCTCCAAACCGTTGATCGAAACCCCTCAAACTCTTCGAG	272	
Db	193	AGCTTCGCACACTTTATGATTATCTTTGCCATTGATGAGTGTGATGAATCTCTTCAG	252	
QY	273	AGAACGAGAGGTAATTCGTTGGATTCAGCTTCGGAGGCATCAACATCGCTCTCGCG	332	
Db	253	CAGATGAGAAAGTTATATAGTGGGGCATAGTCTTGGTGGTATGAATTTGGGACTTGCTA	312	
QY	333	CGACATATTTCCGGCGAAGATTAAGGTTCTTGTCTCTCAACGCTCTCTGCCCGACA	392	
Db	313	TGMAAAGTATCCAAAGATCTATGCTGCTGTTTTCTTGCTGCTTTCATGCTGATT	372	
QY	393	CAACCCAGCTGCTTCTACGTTCTGGCAAGTATATGAGAGATGCTGGAG-----GTT	446	
Db	373	CTGTTCAACAACCTCTCTTGTGTTGGAAAGTATAAATGAGCGGACGCCAGCGAGAAT	432	
QY	447	TGGGAGATTGTGATTTTTCATCTCATGAACAAAGAAATGGGACGATGAGTTTATTGAAGA	506	
Db	433	GTTGGATACTCAGTTTTTACCATATGTTTCCCTGGAAGAGCCACTGACATCCATGTTTT	492	
QY	507	TGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGCCATAGAGATTACGAGC	566	
Db	493	TTGGCCCAAGTCTTGGCTCAAGCTCTACCAGTATGCTCTCTCGAGGATCTTGCAT	552	
QY	567	TGGCAAAATTTGTCATAGGCAAGGTCATTTTTCACAGAGATCTATCAAGAAGAAA	626	
Db	553	TAGCATCATCAVTTGGTGAGCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAGT	612	
QY	627	AGTTAGCAGGAAGATATGTTTCGTCACAGGATTTACCTAATGAGTAGTGAAGACA	686	
Db	613	ATTTACAGATGAACGGTTGGATCAGTGAAGAGAGTTTACATTTGTGCTAGGAGATA	672	
QY	687	AAGCCATCCCTCGCATTTCAVTTGTTGATGATTGATAATTTCAACGCTCTCGAAAGTCT	746	
Db	673	AAGCATACAGAGAAATTCACGCATGGAATTTGACAAATTTGGTGTCTACATGAAGCNA	732	
QY	747	ACGAGATCGATCGGAGATCAATGGTATGCTCTTCCAAACCCCAAAAACCTCTTTGACT	806	
Db	733	TAGAGATTAAAGTGCTGATCAGTGGCAATGCTATGCGAGCCCAAAAACCTTTGCGCT	792	
QY	807	CTCTCTCTGCTATTCGCCCGATTATATGTAAT	839	
Db	793	CTCTCTTGGAAATTCGCCCAATTAATCAACTGAT	825	

LOCUS	HEU40402	1078 bp	mRNA	linear	PLN 13-MAR-1996
DEFINITION	Hevea brasiliensis hydroxynitrile lyase (hnl) mRNA, complete cds.				
ACCESSION	U40402				
VERSION	U40402.1	GI:1223883			
KEYWORDS					
SOURCE	Hevea brasiliensis (Para rubber tree)				
ORGANISM	Hevea brasiliensis				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids 1; Malpighiales; Euphorbiaceae; Crotonoideae; Micrandreae; Hevea.				
REFERENCE	1 (bases 1 to 1078)				
AUTHORS	Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S.D. and Schwab, H.				
TITLE	Molecular cloning of the full-length cDNA of (S)-hydroxynitrile lyase from Hevea brasiliensis. Functional expression in Escherichia coli and Saccharomyces cerevisiae and identification of an active site residue				
JOURNAL	J. Biol. Chem. 271 (10), 5884-5891 (1996)				
MEDLINE	96215058				
PUBMED	8621461				
REFERENCE	2 (bases 1 to 1078)				
AUTHORS	Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S.D. and Schwab, H.				
TITLE	Direct Submision				
JOURNAL	Submitted (09-NOV-1995) Meinhard Hasslacher, Department of Biochemistry, University of Technology Graz, Petersgasse 12/2, Graz, 8010, Austria				
FEATURES	Location/Qualifiers				
source	1..1078				
	/organism="Hevea brasiliensis"				
	/mol_type="mRNA"				
	/db_xref="taxon:3981"				
gene	1..1078				
	/tissue_type="leaves"				
CDS	44..817				
	/gene="hnl"				
	/gene="hnl"				
	/codon_start=1				
	/product="hydroxynitrile lyase"				
	/protein_id="AAC49184.1"				
	/db_xref="GI:1223884"				
	/translation="MAFAHFLVLIHTICHGAWIWHKLPLEALGHKVTALDLAASGVD PROLEIGSDPEYSEPLLTLEALPGEKVLAVGESGGLNIAIAADKVCCKIAAAV HNSVDPTEHCPSTVVDKLMVEVPDWDKDTYFTYTKGKEITGLKFTLLRENLYTL CGPEYBLAKMLTRKSLFQNLAKRPFFTKBEGYSIRKIIYVTDQDEIFLPEFQLMQ IENYKPKDKVYKVEGGDHKLQLTKTKEIAEILQEVADTYN"				
ORIGIN					
Query Match	24.3%;	Score 221.6;	DB 8;	Length 1078;	
Best Local Similarity	55.3%;	Pred. No. 8.2e-52;			
Matches 431;	Conservative 0;	Mismatches 349;	Indels 0;	Gaps 0;	
QY	76	TCACCTTCGTGTAGTTTCAACACGCTTATCATCGAGCCCTGGATCTGGTCAACAGCTCAAGCC	135		
Db	55	TCATTTTGTCTTATTCATACCATATGCCAGGTGCATGGATTTGGCACAGCTCAAAACC	114		
QY	136	CCTCTTGAATCAGCCGCCACCGGTTACTGTGTGCAACTCGCCGCTCCGGATCGA	195		
Db	115	CCTCTTGGAGCACTTGGCCACAAGGTACTCTGCATCGACCTTGACGCAAGCGCGTTGA	174		
QY	196	CCACGACCATCAGGCCGTTGAAACCGTCGACGAACTACTCCAAACCGTTGTATCGAAGC	255		
Db	175	CCCAAGGCAATGAGGAGATTGGCTCATTTGATGAGTATTTCTGAACCTTTGTGACGTT	234		
QY	256	CCTCAAACTCTTCCAGAGAACGAGGTAAATCTTGGTTGGATTTCACTTCGGAGGCAAT	315		
Db	235	CTTGGAGCACTCTCCCTCCAGGGGAAAGGTGATTTCTGTTGCGGAGAGCTGTGGAGACT	294		
QY	316	CAACATCGCTCTCGCCGCCGACATATTTCCCGCGGAAGATTAAAGTTCTTGTGTTCTCAA	375		
Db	295	CAATATAGCAATTTGCTGCTGATAAATACTGTGAAAAGATTGACGTGCTGTTTTCACAA	354		

QY 376 CGCTTCTTGGCCGACACCAACCCACGTCGCTTCTCAAGTTCTGGACAAGATATATGAGAT 435  
Db 355 TTCAAGTATTGCCAGACACCGAGCAGTCCCATCTTACGTCGTGATAAAGCTCATGAGGT 414  
QY 436 GCCTGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGGGACGATGAG 495  
Db 415 GTTTCGCGACTGGAAAGACACCAAGTATTATTCGTACACTAAAGATGGCAAGGAGATAAC 474  
QY 496 TTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGTCCCATAGA 555  
Db 475 TGGATTGAAGTGGGCTTACGCTTCTGAGGGAATTTATATACCTTTGGGCTCTGA 534  
QY 556 GGATTACGAGCTGGCAAAATTTGTCATAGCGCAAGGTCATTTTTCACAGAGATCTATC 615  
Db 535 GGAATATGAACCTGGCAAGATTTGACAAGGAAGGATCATTTTCAAAATATTTTATG 594  
QY 616 AAGAAGAAAGATTTAGCGGAAGGATATGTTGCGTGCAACGATTTAGTAATGAG 675  
Db 595 TAAGCGACCAATCTTCACTAAGGAAGGTTACGGATCGATTAAGAAAATTTATGTGGAC 654  
QY 676 TAGTGAAGACAAAGCCATCCCTCGGATTTTCACTTCTGATGATGATTAATTTCAACGT 735  
Db 655 CGACCAAGACGAATATTTTACCTGATTTCACTCTGCGCAATAGAAAATCTATAAAC 714  
QY 736 CTCGAAAGTCTACGAGATCGATGGCGGAGATCATGTTGATGCTCTCCAAACCCCAAA 795  
Db 715 AGACAAGTTTATAAGTTCGAAGTGGAGATCATATAAATTCAGCTTACAAAGACTAAGGA 774  
QY 796 ACTCTTTGACTCTCTCTGCTATTCGCCACGATTTATATGTAATATCTTAAGTCGGTTT 855  
Db 775 GATCGCTGAATTTCTCAAGAGGTGGCTGATACCTATATAATGACTCTTTTGAGGCTTTT 834

RESULT 14  
A59586  
LOCUS A59586 1091 bp DNA linear PAT 06-MAR-1998  
DEFINITION Sequence 1 from Patent WO9703204.  
ACCESSION A59586  
VERSION A59586.1 GI:3714897  
KEYWORDS Hevea brasiliensis (Para rubber tree)  
SOURCE Hevea brasiliensis  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;  
Micrandreae; Hevea.  
REFERENCE 1  
AUTHORS Hasselacher, M., Schall, M., Schwab, H., Hayn, E.M., Kohlwein, S. and Griengl, H.  
TITLE (S)-HYDROXYNITRILLYASE FROM HEVEA BRASILIENSIS  
JOURNAL Patent: WO 9703204-A 1 30-JAN-1997;  
DSM CHEMIE LINZ GMBH (AT)  
FEATURES  
source  
1. .1091  
/organism="Hevea brasiliensis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3981"  
gene 1. .1091  
/gene="HNL"  
mRNA 1. .1091  
/gene="HNL"  
/function="(S)-HYDROXYNITRILLYASE"

ORIGIN  
Query Match 24.3%; Score 221.6; DB 6; Length 1091;  
Best Local Similarity 55.3%; Pred. No. 8.2e-52;  
Matches 431; Conservative 0; Mismatches 349; Indels 0; Gaps 0;  
QY 76 TCACCTCGTGTAGTTCAACGCTTATCATGAGCTGGATCTGTACAAGCTCAAGCC 135  
Db 68 TCATTTTGTCTTATTCATACCATATGCCCAGCTGCATGATTTGGCAAGCTCAAAAC 127  
QY 136 CCTCCTTGAATCAGCGCGCCACCGCTTACTGTGTGAACTCGCGGCTCGGGATCGA 195

Db 128 CCTCCTTGAAGCACTTGGCCACAAGGTTACTGCACTGGACCTTGCAGCAAGCGCGCTTGA 187  
QY 196 CCCAGCACAATCCAGGCGCTTGAACCGTCCAGCAATCTCCAAACCGCTTGCAGAAC 255  
Db 189 CCCAAGGCAATTTGAGGAGATTGGCTCATTTGATGAGTATTTCTGAACCCCTTTGACGTT 247  
QY 256 CCTCAAAATCTCTTCCAGAGAACGAAGAGATTAATCTGGTTGGATTCAGCTTCGGAGGCAT 315  
Db 248 CTTGAGGCACTCCCTCCAGGGGAAGGTGATTCTGGTTGGGAGAGCTGTGGAGACT 307  
QY 316 CAACATCGCTCTCGCGCGCACATATTTCCGGCGAAGATTAAGGTTCTTGTCTCTCAA 375  
Db 308 CAATATAGCAATGTCTGTGATAAATACTGTGAAAAGATTGACGTGCTGTTTTCACAA 367  
QY 376 CGCTCTTCTGCGCGCACAAACCCAGTCCTCTCACGTTCTGGCAACAGTATATCGAGAT 435  
Db 368 TTCAGTATTGCCAGACACCGACACTGCCCATCTTACGTCGTGGATAAGCTCATGAGGT 427  
QY 436 GCCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAAACAAGAAATGGGACGATGAG 495  
Db 428 GTTTCGCACTGGAAAGACACCACTATTTTACGTACACTAAGATGGCAAGAGATAAC 487  
QY 496 TTTATTGAAGTGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCATAGA 555  
Db 488 TGGATTGAACCTGGGCTTCACTCTCTGAGGGAATTTATATACCTTTGGGCTCTGA 547  
QY 556 GGATTACGAGTGGCAAAATTTGTCATAGCAAGGTCATTTTTCACAGAGATCTATC 615  
Db 548 GGAATATGAACCTGGGGAAGATGTTGACAAAGGAAGGATCATTTTTCAAAATATTTTACG 607  
QY 616 AAGAAGAAAGATTTTAGCGGAAGGATATGTTTGGTGGTCAACGAGTTTACGTAATGAG 675  
Db 608 TAAGCGACCATCTTCACTAAGAGGTTACGATCGATTAGAATAATTTATGTGGAC 667  
QY 676 TAGTGAAGACAAAGCCATCCCTCGGATTTCACTGTTGATGATGTAATTTCAACGT 735  
Db 668 CGACCAAGACGAAATATTTTACCTGAAATTTCACTCTGGCAATAGAAAACATATAAAC 727  
QY 736 CTCGAAAGTCTACGAGATCGATGGCGGAGATCAGATGATGCTCTCCAAACCCCAAA 795  
Db 728 AGACAAGGTTTATAAGGTCGAAGTGGAGATCATAAATTTGACGCTTACAAAGACTAAGA 787  
QY 796 ACTCTTTGACTCTCTCTGCTATTTGCCACCGATTTATATGTAATATCTTAAGTCGGTTT 855  
Db 788 GATCGCTGAATTTCTTCAAGAGGTGGCTGATACCTATTAATTCACCTCTTTGAGGCTTTT 847

RESULT 15  
CQ806104  
LOCUS CQ806104 792 bp DNA linear PAT 10-MAY-2004  
DEFINITION Sequence 2515 from Patent WO2004035798.  
ACCESSION CQ806104  
VERSION CQ806104.1 GI:47111686  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1  
Inze, D., de Veylder, L. and Vlieghe, K.  
IDENTIFICATION of novel e2f target genes and use thereof  
Patent: WO 2004035798-A 2515 29-APR-2004,  
CropDesign N.V. (BE)  
FEATURES  
Location/Qualifiers  
source 1. .792  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 24.0%; Score 218.8; DB 6; Length 792;





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 21:55:38 ; Search time 609 Seconds  
(without alignments)  
2803.748 Million cell updates/sec

Title: US-10-780-002-2  
Perfect score: 1368  
Sequence: 1 MKEGKHFVLVHGAGCHGWSW.....MLCEPQKLCASLEIAHKYN 260

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USFTO.spool\_p/US10780002/runat\_11092005\_104749\_26712/app query.fasta\_1.455  
-DB=Published Applications NA -QWTS=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10780002 @CGN 1 1 723 @runat\_11092005\_104749\_26712  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1368	100.0	1079	21	US-10-780-002-1	Sequence 1, Appli
2	837	61.2	933	18	US-10-424-599-64748	Sequence 64748, A
3	768	56.1	792	9	US-09-938-842A-1320	Sequence 1320, Ap
4	768	56.1	792	11	US-09-938-842A-1320	Sequence 1320, Ap
5	768	56.1	989	21	US-10-780-002-32	Sequence 32, Appl
6	742	54.2	1158	18	US-10-424-599-85023	Sequence 85023, A
7	736	53.8	1035	18	US-10-425-114-12800	Sequence 12800, A
8	695	50.8	792	9	US-09-938-842A-1551	Sequence 1551, Ap
9	695	50.8	792	11	US-09-938-842A-1551	Sequence 1551, Ap
10	695	50.8	870	21	US-10-780-002-33	Sequence 33, Appl
11	690	50.4	792	9	US-09-938-842A-217	Sequence 217, App
12	690	50.4	792	11	US-09-938-842A-217	Sequence 217, App
13	690	50.4	837	21	US-10-780-002-34	Sequence 34, Appl
14	680.5	49.7	792	21	US-10-780-002-35	Sequence 35, Appl
15	669.5	48.9	1101	19	US-10-437-963-62048	Sequence 62048, A
16	665.5	48.6	819	21	US-10-780-002-37	Sequence 37, Appl
17	659	48.2	1069	19	US-10-767-701-9317	Sequence 9317, Ap
18	654.5	47.8	834	21	US-10-780-002-39	Sequence 39, Appl
19	645	47.1	783	21	US-10-780-002-38	Sequence 38, Appl
20	640.5	46.8	771	9	US-09-938-842A-1278	Sequence 1278, Ap
21	640.5	46.8	771	11	US-09-938-842A-1278	Sequence 1278, Ap
22	640.5	46.8	771	21	US-10-780-002-40	Sequence 40, Appl
23	635	46.4	912	21	US-10-780-002-36	Sequence 36, Appl
24	616	45.0	653	18	US-10-424-599-52897	Sequence 52897, A
25	613	44.8	1007	18	US-10-425-114-21611	Sequence 21611, A
26	613	44.8	1018	20	US-10-425-115-119846	Sequence 119846, A
27	600	43.9	1025	18	US-10-425-114-13774	Sequence 13774, A
28	600	43.9	1092	20	US-10-425-115-70617	Sequence 70617, A
29	590	43.1	901	21	US-10-780-002-41	Sequence 41, Appl
30	578	42.3	996	18	US-10-767-701-15339	Sequence 15339, A
31	555.5	40.6	907	18	US-10-425-114-24341	Sequence 24341, A
32	555.5	40.6	921	20	US-10-425-115-70618	Sequence 70618, A
33	550	40.2	1039	18	US-10-425-114-24225	Sequence 24225, A
34	550	40.2	1249	20	US-10-425-115-179879	Sequence 179879, A
35	548	40.1	1087	19	US-10-437-963-23964	Sequence 23964, A
36	548	40.1	1148	19	US-10-437-963-29636	Sequence 29636, A
37	545	39.8	786	19	US-10-437-963-62697	Sequence 62697, A
38	545	39.8	1173	19	US-10-437-963-89948	Sequence 89948, A
39	538.5	39.4	858	19	US-10-437-963-75962	Sequence 75962, A
40	530.5	38.8	777	21	US-10-738-922-1	Sequence 1, Appli
41	530.5	38.8	777	21	US-10-738-922-2	Sequence 2, Appli
42	530.5	38.8	777	21	US-10-738-927-1	Sequence 1, Appli
43	530.5	38.8	777	21	US-10-738-927-2	Sequence 2, Appli
44	527.5	38.6	920	18	US-10-424-599-128189	Sequence 128189, A
45	526	38.5	1006	20	US-10-425-115-36145	Sequence 36145, A

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klesig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE REFERENCE: 3670-PO2652W0  
; CURRENT APPLICATION NUMBER: US/10/780.002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16

```
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 1079
/ TYPE: DNA
/ ORGANISM: Nicotiana tabacum
US-10-780-002-1

Alignment Scores:
Pred. No.:      7,15e-164      Length:      1079
Score:          1368.00      Matches:      260
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              21          Gaps:          0

US-10-780-002-2 (1-260) x US-10-780-002-1 (1-1079)

QY 1 MetLysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrp 20
DB 42 ATGAAGGAAGGAAACACTTTGTTTGTAGTACATGCTGTCATGCGATGCGAGGTTGGAGTTGG 101
QY 21 TyrLysLeuLysProLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 40
DB 102 TACAAGCTAAAGCCACTGCTAGAGCTGAGGCCCATAGGTTACAGCCCTTGATTTAGCA 161
QY 41 AlaSerGlyThrAspLeuArgLysIleGluGluLeuAlaArgThrLeuTyrAspTyrThrLeu 60
DB 162 GTCTTGGCAGCTGATTTGAGAAAATAGAGAGCTTCGCACACTTTATATATATCTTTG 221
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 222 CCATTGATGAGCTGATGAAATCTCTTTGAGCAGATGAGAGGTTATATTTAGTGGGCGAT 281
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 282 ATGCTTGGTGGTATGAATTTGGGACTTGTATGGAAAAATATCCAAAAAGATCTATGCT 341
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
DB 342 GCTGTTTCTTGGCTGCTTTCATGCTGATCTGTTTCACTCCCTCTCTTGTGTTGGAA 401
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
DB 402 CAGTATAATAGCGGACGCGCCAGCAAGATGCTGGTGGATCTCAGTTTATACCATATGCT 461
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
DB 462 TCCCTTGAAGAGCCACTGACATCCATGTTTGTGCCCCAAGTTTCTTGCTCACAAGCTC 521
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
DB 522 TACCAGCTATGCTCTCTCTGAGGATCTTGCAATTAGCATCATCTTGGTGAGACCAAGCTCT 581
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
DB 582 TTGTTTATGGAAGACCTATCGAAGGCCAAGTATTTACAGATGAACGGTTTGGATCAGTG 641
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
DB 642 AAGAGAGTTTACATTGTGTGCACTGAGGATAAAGGCATACCAGAGAAGATTTCCAGCGATGG 701
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyValaAspHisMetAla 240
DB 702 CAAATTGACAACTTGGTGTCTACTGAAGCAATAGAGATTAAAGGTGCTGATCACATGCA 761
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260
DB 762 ATGCTATGCGAGCCCCCAAAACTTTTGGCGCTCTCTCTTTGGAAATTTGCCATAAATACAC 821

RESULT 2
US-10-424-599-64748
; Sequence 64748, Application US/10424599

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 64748
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29480C.1
US-10-424-599-64748

Alignment Scores:
Pred. No.:      3.45e-96      Length:      933
Score:          837.00      Matches:      156
Percent Similarity: 76.86%      Conservative: 40
Best Local Similarity: 61.18%      Mismatches: 55
Query Match:      61.18%      Indels:      4
DB:              18          Gaps:          1

US-10-780-002-2 (1-260) x US-10-424-599-64748 (1-933)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
DB 43 AAGCACTTTGTTGTTGTCATGGGCGATGCCATGGAGCATGGTGTGTGTATTAAGCTCAAG 102
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
DB 103 CCAGCTTGGGAATCTGCTGGGCACAGGTACAGTCTTGACCTTGACCTTGGAGGCC 162
QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
DB 163 AACATGAAGAAAATTTGAAGATGTTGACACTTTTTTCACAGTATACTGAGCCCTTTGTTGTTT 222
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
DB 223 CTATTGGACACAATTCCTCAATGAAAGAGTAGTCTTAGTGTGTCACAGCTTTGGAGGG 282
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
DB 283 CTGAACATAGACACTTGCATGGAGAAATTTCCAGAAAAGGTAGCAGTTGGTGTGTTTCTA 342
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
DB 343 ACAGCTTTGCTCCACAGCTTGAACACACCACCATCTTATGTCTTGGAAAAGTACAGCGAG 402
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
DB 403 AGGACCCCGTTAGCTGCACTGCTTACACTGAATTTGCTCCAGTGGAAAAA----- 456
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
DB 457 -----ACACAAATGTTCTTTGGCCCCCAACTTCTTGTCGCAAGCTCTACCAACTATCC 510
QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
DB 511 CCAATTGAGGATTTGAAATGGCCCAAGACTTTAATAAGGCCATCATCTCTTTCATGGAA 570
QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
DB 571 GACTTGATCAACAAAAGAACTTTCTCCAAAGAGGATATGGGTGAGTCCCGCCGCTCTT 630
QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
DB 631 ATTTGTTGCACTGAGGACCTTGCAATTCATTTGGAATATATCAGCTCTTTCATGATCCAAAT 690
```

Oy	225	IleGlyValThrGluaIalleDlulleIysGlyAlaIalepHisMetalMetLeuCysGlu	244
		::::	
Db	691	GTTGGGTCAATGAAGTCTTAGAGATCAAAGCGCAGACCATATGGCTATGCTTTCCAAG	750
		::::	
Oy	245	ProGlnLysLeuCysAlaSerLeuLeuGluAlaHisLysTyr	259
		::::	
Db	751	CCCCAAGAACAATTGATTCCTCCAGCAGCATAGCACATAAAT	795
		::::	

```

RESULT 3
US-09-938-842A-1320
; Sequence 1320, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1320
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1320

```

Alignment Scores:		
Pred. No.:	1.73e-87	Length:
Score:	768.00	Matches:
Percent Similarity:	71.04%	Conservative:
Best Local Similarity:	57.53%	Mismatches:
Query Match:	56.14%	Indels:
DB:	9	Gaps:
		2
		2
US-10-780-003-2 (1-260)	x	US-09-938-842A-1320 (1-792)

Qy	2	LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr	21
Db	13	AAGAGAGAAAACAACATTTTGTACTAGTACATGGTTCGTGCCATGGCGGTGTCGTGGTAC	72
Qy	22	LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaA	41
Db	73	AMGTTTAAGCCGCTGCTAGAGCGGTGGCCACCGCTAACTGCTGGAGCTTAGTGTGCC	132
Qy	42	SerGlyThrAsp---LeuArgLysIleGluGluArgThrLeuTyrAspTyrThrLeu	60
Db	133	TCCGGAATAGACAACAACGAGGTGCGATCACTGACATCCCCACATCGCAACAATCTCGGAG	192
Qy	61	ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis	80
Db	193	CCATTGACGAAGCTCTCCACCTCATTTGCCAAATGATGAAAGGTTGTGCTGTGGTGTAC	252
Qy	81	SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla	100
Db	253	AGCTTTGGTGGCTTTGAACATTAGCCATAGCCATGAAAGATTTCCCGAAAAATCTCTGTCT	312
Qy	101	AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu	120
Db	313	GCTGTATTCTTGACTGCTTTCATGCGGACACCGCACTCACCATCTTCGTCTTGGAC	372
Qy	121	GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly	140
Db	373	AAAGTTTGAAGACAACATGCTCAAGAAGCATGTGATGGGCACCGAAATTCGAACCTTATGCT	432

Qy	141	SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu	160
Db	433	TCAGACAATTCGGAGCTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGCTCTC	489
Qy	161	TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer	180
Db	490	TACCAGCTTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATGAGCCGAGATCG	549
Qy	181	LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal	200
Db	550	TTATTTTAAACGATTTATCGAGATGAAAACTCTCGGATGAGGATATGGGTCGTGT	609
Qy	201	LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr	220
Db	610	CCTCGAGTTTTTCATGTGTGTAAAGAGGACAAAGCAATTTCCAGAGAAACGCCAGAGATGG	669
Qy	221	GlnIleAspAsnIleGlyValThrGluAlaIleGluIleGlyGlyAlaAspHisMetAla	240
Db	670	ATGATTGATAATTTTCGGGTGAATTTTAGTGATGGAGATGGAGGAGACAGATCATATGCCA	729
Qy	241	MetLeuCysGluProGlnIlyLysLeuCysAlaSerSerLeuLeuGluIleAlaHisLysTyr	259
Db	730	ATGTTCTCGAAGCTCTCAGAACTTCAGTGATTTACTTCTGAAATTCGCGACAAATTC	786

RESULT 4  
US-09-938-842A-1320  
; Sequence 1320, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1320  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1320

```

Alignment Scores:
Pred. No.:      1,73e-87          Length:         792
Score:          768.00           Matches:        149
Percent Similarity:    71.04%       Conservative:   35
Best Local Similarity: 57.53%       Mismatches:     73
Query Match:         56.14%        Indels:         2
DB:                11              Gaps:          2

US-10-780-002-2 (1-260) x US-09-938-842A-1320 (1-792)

Qy      2 LysGluGlyHisSpheValLeuAlaGlyHisGlyAlaCyeHisGlyGlyTrpSerThrTyr 21
      ||| :|| ::|||::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      13 AAGAGGAACAACATTTTGCTACTAGTACATGGTTTCGTGCATGCCGCGTGCTGGTAC 72

Qy      22 LysLeuLysProteinLeuGluaAlaGlyHISlysValThrAlaLeuAspLeuAlaLa 41
      ||| :|| ::|||::| |::| |::| |::| |::| |::| |::| |::|
Db      73 AAGTTTAGCCGCTGCTAGAAGCGGTGGGCGCACCGGCTAACTGCTGTGACCTAGCTGCC 132

Qy      42 SerGlyThrASP---LeuArgLysIleGluGluLeuAurgThrLeuTyrAspTyrThrLeu 60
      ||| :|| ::|||::| |::| |::| |::| |::| |::| |::| |::|
Db      133 TC CGGAATAGACAACAACGAGGTGCATCTCATCTCAATCCCCACATCGCAACAATACTCGGAG 192

```

```
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 193 CCATTGACGAGCTCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTGCTGGTCAC 252
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 253 AGCTTTGGTGGCTTGAACATTAGCCATAGCCATGGAAAGTTTCCCGAAAAAATCTCTGTC 312
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 313 GCTGTATTCTTCACTGCTTTCATGCGGACACCGAACACTCACCATCTCTGCTTGGAC 372
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 373 AAGTTTGGGAAGCAACATGCTTCAAGAAGCATGGATGGGCACCGAATTCGAACCTTATGCT 432
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 433 TCAGACAATTCGGAGCTG---AGTATGTTTGTAGCCCTGACTTCATGAAGTTGGGTCTC 489
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 490 TACCAGCTTCTCCAGTTGAGGATCTTGAACTGGGATTACTTTTATGAGGCCAGGATCG 549
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 550 TTATTATTAAAGCATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGTT 609
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 610 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTTCCAGAGAAGCCGACAGATGG 669
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 670 ATGATTGATAATTTCCGGTGAATTAGTGATGGAGATGGAGGACAGATCATATGCCA 729
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 730 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAATCGCGACAAATTC 786

RESULT 5
US-10-780-002-32
; Sequence 32, Application US/107800002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: 3670-PO2652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 989
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-32
```

```
Alignment Scores:
Pred. No.: 2,42e-87 Length: 989
Score: 768.00 Matches: 149
Percent Similarity: 71.04% Conservative: 35
Best Local Similarity: 57.53% Mismatches: 73
Query Match: 56.14% Indels: 2
DB: 21 Gaps: 2
```

```
US-10-780-002-2 (1-260) x US-10-780-002-32 (1-989)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 119 AAGAGGAACAACATTTTGTACTAGTACATGGTTCTGTCGCCATGGCGGCTGCTGTGGTAC 178
QY 22 LysLeuLysProLeuLeuGluAlaGlyHisValThrAlaLeuAspLeuAlaAla 41
Db 179 AAGGTTAAGCGCTGCTAGAGCGGTGGGCCACCGCTAACTGCTGTAGCTTATAGCTGCC 238
QY 42 SerGlyThrAsp---LeuArgLysIleGluLysLeuArgThrLeuTyrAspTyrThrLeu 60
Db 239 TCCGGATAGACACAACGAGGTGCTACTGACATCCCAACATCGAACAATATCTCGAG 298
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 299 CCATTGACGAAGCTCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTGCTGGTCAC 358
QY 81 SerLeuGlyGlyMetAsnLeuLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 359 AGCTTTGGTGGCTTGAACCTTAGCCATAGCCATGAAAGTTTCCCGAAAAAATCTCTGTC 418
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 419 GCTGTATTCTTCACTGCTTTCATGCGGACACCGAACACTCACCATCTCTGCTTGGAC 478
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 479 AAGTTTGGGAAGCAACATGCTTCAAGAAGCATGGATGGGCACCGAATTCGAACCTTATGCT 538
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 539 TCAGACAATTCGGAGCTG---AGTATGTTTGTAGCCCTGACTTCATGAAGTTGGGTCTC 595
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 596 TACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATGAGGCCAGGATCG 655
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 656 TTATTATTAAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGTT 715
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 716 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTTCCAGAGAAGCCGACAGATGG 775
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 776 ATGATTGATAATTTCCGGTGAATTAGTGATGGAGATGGAGGACAGATCATATGCCA 835
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 836 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAATCGCGACAAATTC 892
```

## RESULT 6

```
US-10-424-599-85023
; Sequence 85023, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85023
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Glycine max
```

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_47790C.1  
US-10-424-599-85023

Alignment Scores:  
Pred. No.: 6,4e-84 Length: 1158  
Score: 742.00 Matches: 142  
Percent Similarity: 70.43% Conservative: 39  
Best Local Similarity: 55.25% Mismatches: 72  
Query Match: 54.24% Indels: 4  
DB: 18 Gaps: 1

US-10-780-002-2 (1-260) x US-10-424-599-85023 (1-1158)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21  
DB 193 AGAGACCAAAATCCTTTGCTCTGGTGATGATAGGCATGGTCCCTGGTGGTAC 252  
QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
DB 253 AAGCTTAAGCCACTGTTGGAATCCGCCGCCCAAAAGTCACAGTCCTTGACCTTGCGACT 312  
QY 42 SerGlyThrAspLeuArgLysLysGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
DB 313 TCTGGCATCGACACACACGACATTTGAAGACATCCACACATTTCTGAGTATTTAAGCCT 372  
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValLeuValGlyHisSer 81  
DB 373 TTGTTGGATCTCTGGCGTCTGCTCTTAATGAAAAGTGCTCTTGTGGGCATAGC 432  
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101  
DB 433 TTTGGAGGGATCAGTATAGCCTTGCATGGCAATGGCAAAATCCAGAGAAATATACCTTGA 492  
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121  
DB 493 ATTTCTCTAACAGCTTTTGTCTGTATACCAACCAACCATCATGCTCTTAGAAGAG 552  
QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141  
DB 553 TACATTGATGATACCATATACCGGATGGATGGACACTGAGCTCTGGAAATAGTGAGGC 612  
QY 142 ProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyr 161  
DB 613 AAA-----ACAACATTGCTTTTGGCATCAAATTTCTGCCACTAAGTCTAT 660  
QY 162 GlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeu 181  
DB 661 CAACCTCTGCCACTGAGGATCTGGAATTTGGTGAAGACTTTAAGAGAAAGGGTTCACTA 720  
QY 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201  
DB 721 TTGCTGAAGACCTTCTAAGCGAGAAATTTTCCAAAGAGAAAGATGGGTCTGTCCA 780  
QY 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGln 221  
DB 781 AGTGCTTATATTATTTCCAATGAGGACTTGGTAATTTCCAAGGAGTATCAGCAATGGATG 840  
QY 222 IleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMet 241  
DB 841 ATCCAAAATGCGAGGATTTGATGTGGTGGCAGAGATCAAGGGATCAGATCAGATGGTATG 900  
QY 242 LeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLys 258  
DB 901 CTTAGCAAAACCCACAAACTATGTTTATCTCTCTCGAGATAGCTGATAAG 951

RESULT 7

US-10-425-114-12800  
Sequence 12800, Application US/10425114  
Publication No. US2004003488A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 12800  
LENGTH: 1035  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: 701212327\_FLI  
US-10-425-114-12800

Alignment Scores:  
Pred. No.: 3,15e-83 Length: 1035  
Score: 736.00 Matches: 143  
Percent Similarity: 72.66% Conservative: 43  
Best Local Similarity: 55.86% Mismatches: 62  
Query Match: 53.80% Indels: 8  
DB: 18 Gaps: 3

US-10-780-002-2 (1-260) x US-10-425-114-12800 (1-1035)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24  
DB 105 AAGCATTATTTCTGTCATGGAGCATGCCACGGAGCTTGGATTGGTGCAAGCTGAAG 164  
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44  
DB 145 CCACCTTGGAAATCGAAGGCCACCAAGGTACAGTGTGCAGCTTGCAGCTTCTGSCATC 224  
QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64  
DB 225 AACATGAAGAAATTCAGATGTTGCACACTTCTCGCAGTACTCTGAGCCTTTGTTCGAG 284  
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84  
DB 285 CTAATGACCAAAATTCCTCAATGAGAAGGTGGTCTTAGTTGGTCACAGTTTGGAGGG 344  
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104  
DB 345 ATGAACATAGCCCTTGCTATGGAGAAATTCAGAAAAGGTTCGAGTGGTGGTCTTTCTTA 404  
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124  
DB 405 ACAGCTTTTGTCTCCAGACACCCGACATCGCCCATCTTACGCTTTGGACAG----- 455  
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144  
DB 456 AACAGCTCGCTCTGAAGAC---TTGGACAACGAATTTGCCCAAGTGGAAACAAG----- 506  
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164  
DB 507 -----ACATCAGTGCTCTTTGGTCTGAGTACTTGTCCAAAGCAATATACAGCTCTCT 560  
QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184  
DB 561 CCCGTTGAGGACCTTAGAATTCGCCAAGACATTTGGTAAGGCCATCATCTCTTCATTGNA 620  
QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204  
DB 621 GATTGTCTAAGCAAAAGAACTTCTCCAAACATGATATGGTTCAGTTCCAGTCCAGCTCTAT 680  
QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224  
DB 681 ATTGTTGCACTGAGGACCTTGCATTTGCAATATGATATGATGATGATGATGATGATGAT 740  
QY 225 IleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAlaMetLeuCysGlu 244

Db 741 GCTGGATCAATGACGCTTCTAATGATCAAAAGTGCAGATCATGCGAGCTATGTTTAGCAAG 800  
Qy 245 ProGlnLysLeuCysAlaSerLeuGluAlaHisLysTyrAsn 260  
Db 801 CCTAGAGACTATTCAATCTCTCCAGAGATAGTACAAAATATGAC 848

## RESULT 8

US-09-938-842A-1551  
; Sequence 1551, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1551  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1551  
Alignment Scores:  
Pred. No.: 3 6e-78 Length: 792  
Score: 695.00 Matches: 138  
Percent Similarity: 67.18% Conservative: 36  
Best Local Similarity: 53.28% Mismatches: 83  
Query Match: 50.80% Indels: 2  
DB: 9 Gaps: 2  
US-10-780-002-2 (1-260) x US-09-938-842A-1551 (1-792)

Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21  
Db 13 AAGAGGAAGCAACACTTCGTGTGTAGTACATGCTGGCGCCGCGCATGGTGGTAC 72  
Qy 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
Db 73 AAGGTTAAGCCTCTTCTCAGAGCTTTGGGCCCATCGGTAAACCGCCTTAGACCTAGCTGCT 132

Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
Db 133 TCCGGTATAGACACACACAGGTCATCTAGTACATCTTCTAGATGGAACAATATCTGAG 192  
Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
Db 193 CCATTGATCAGCTAATGACTTTCATTCGCCAATGATGAGAGGTTGTACTCGTTGGTCAT 252  
Qy 81 SerLeuGlyClyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
Db 253 AGCTTTGGAGGTTTGAGTTTAGCCTTAGCCATGGATAGTTTCCCGGATAAATCTCTGTC 312

Qy 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
Db 313 TCTGCTTCGTGACTGCTATTCATGCCGACACAAACACTACCATCGTTCGTCGAGNA 372  
Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
Db 373 AAGTTTGCAAGCAGCATGACACCAAGAGGATGGATGGGGCTCTGAGCTCGAGACATATGT 432  
Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160

Db 433 TCAGATAATTCGGGCTTG---TCTGTGTTCTTCAGACCAGCTTCATGACACCGCTC 489  
Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
Db 490 TACCAACTTTCTCCTGTGGAGGATCTTGAGCTTGATTTGCTTCTAAAGAGGCTAGTTCA 549  
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
Db 550 TTGTTTATTAATGAATATATCGAAGATGGAGAACTTTTCTGAGAAAGGGTATGGATCTGTT 609  
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrp 220  
Db 610 CTCGAGCTTACATGTGTGCAAGAGGACCAACATTATCTCGGAAGACCATCAACGATGG 669  
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240  
Db 670 ATGATCCATAATTTATCCGGCGAATTTAGTGATTGAGATGGAAGAGACTGATCATATGCCA 729  
Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
Db 730 ATGTTTTCGAAAACCTCACTACTAAGTGACCATCTATTGGCAATCGCTGCACAATTTC 786

## RESULT 9

US-09-938-842A-1551  
; Sequence 1551, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1551  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1551  
Alignment Scores:  
Pred. No.: 3 6e-78 Length: 792  
Score: 695.00 Matches: 138  
Percent Similarity: 67.18% Conservative: 36  
Best Local Similarity: 53.28% Mismatches: 83  
Query Match: 50.80% Indels: 2  
DB: 11 Gaps: 2  
US-10-780-002-2 (1-260) x US-09-938-842A-1551 (1-792)

Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21  
Db 13 AAGAGGAAGCAACACTTCGTGTGTAGTACATGCTGGCGCCGCGCATGGTGGTGGTAC 72  
Qy 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
Db 73 AAGGTTAAGCCTCTTCTCAGAGCTTTGGGCCCATCGGTAAACCGCCTTAGACCTAGCTGCT 132  
Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
Db 133 TCCGGTATAGACACACACAGGTCATCTAGTACATCTTCTAGATGGAACAATATCTGAG 192  
Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80

```
Db 193 CCATTGATGACGTAATACTTCAATTCGCCGAATGATGAGAAGGTTGTACTCGTGGTGCAT 252
Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 253 AGCTTTGGAGGTTTGGTTTGGCTTAGCCATGATGAAGTTTCCCGATAAATCTCTGTC 312
Qy 101 AlaValPheLeuAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 313 TCTGTCTTCGTGACTGCAATTCATGCCGACACCAACACTCCACATCGTTGTCGAGGAA 372
Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGly 140
Db 373 AAGTTTTCGAAGCAGCATGACACCAAGGATGGCTCTGAGCTCGAGACATATGTT 432
Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 433 TCAGATAATTCGGCTTG---TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCTCTC 489
Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 490 TACCAACTTCTCCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTACTTCA 549
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 550 TTGTTTATTAAATGAATATCGAAGATGGAGAACTTTTCTGAGAAAGGATATGATCTGTT 609
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr 220
Db 610 CCTCGAGCTTACATTTGTGCAAGAGGACCAACATTATCTCGAAGACCATCAACGATGG 669
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 670 ATGATCCATAATATTCGGCGAATTTAGTCAATTGAGATGGAAGAGACTGATCATATGCCA 729
Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 730 ATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCTGTCACAATTC 786
```

## RESULT 10

```
US-10-780-002-33
; Sequence 33, Application US/10780002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: 3670-PO2652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-33
```

```
Alignment Scores:
Pred. No.: 4.15e-78 Length: 870
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 21 Gaps: 2
```

US-10-780-002-2 (1-260) x US-10-780-002-33 (1-870)

```
Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyr 21
Db 33 AAGAGGAAGCAACACTTCGTGCTAGTACATGTTGGTGGTCCACGGCGCATGGTCTGGTAC 92
Qy 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
Db 93 AAGGTTAAGCCCTTCTTCGAGGCTTTGGGCCATCGGTAAACCGCTTAGACCTAGCTGCT 152
Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 153 TCGGTATAGACACACACAGGTCATCTACATGACATTTCTACATGTAACAATATCTCTGAG 212
Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 213 CCATTGATGACGTAATGACTTCAATTCGCGAATGATGAGAAGTTGTACTCGTTGGTTCAT 272
Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 273 AGCTTTGGAGGTTTGGAGTTTAGCCCTTAGCCATGGATTAAGTTTCCCGATAAATCTCTGTC 332
Qy 101 AlaValPheLeuAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 333 TCTGTTCTCGTCACTGATTCATGCCGACACCAACACTCCATCTGTTCTGTCGAGGAA 392
Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGly 140
Db 393 AAGTTTGCAGCAGCATGACACCAAGGATGGATGGGCTCTGAGCTCGAGACATATGTT 452
Qy 141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160
Db 453 TCAGATAAATTCGGCTTG---TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCGCTCTC 509
Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 510 TACCAACTTCTCCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTACTTCA 569
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 570 TTGTTTATTAAATGAATATTCGAAGATGGAGAACTTTTCTGAGAAAGGGTATGATCTGTT 629
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr 220
Db 630 CCTCGAGCTTACATTTGTGCAAGAGGACCAACATTATCTCGAAGACCATCAACGATGG 689
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 690 ATGATCCATAATATTCGGCGAATTTAGTCAATTGAGATGGAAGAGACTGATCATATGCCA 749
Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 750 ATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCTGTCACAATTC 806
```

## RESULT 11

```
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
```

```
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217

Alignment Scores:
Pred. No.: 1.56e-77 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x US-09-938-842A-217 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrPsrTyrTyrLeuLys 24
DB 22 CAACACGCTGCTTCTAGTACATGCTGCTGCCAGCGGTAGATCTAGCTGCTCGGTATA 141
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
DB 82 CCGAGCTCGAGGCTTCTGGCCACCGGTAAACCGCGGTAGATCTAGCTGCTCGGTATA 141
QY 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
DB 142 GACATGACCGAGTCAATCACAGATATATCCACATCGCAACAATATCTCAGAGCCATTGATG 201
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
DB 202 CAGCTAATGACCTCATTACACGATGATGATGATGATGATGATGATGATGATGATGATG 261
QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
DB 382 AAAGAAACTTCAGAGAGAAATGGTGTACACCGGTGTACAGC---GAGAAACCTGAT 438
QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
DB 439 TTTCCTAGCGAGTTTGGATTCTTGGACCAAGATTCATGGCCAAAGACTTGTATCAGTTG 498
QY 164 CysSerProGluAspLeuAlaSerLeuValArgProSerSerLeuPheMet 183
DB 499 TCTCCAGTCCAAAGATCTTGAATTTGGCGAAATTTGGTGGAGGCGCAACCATTTGATTAG 558
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
DB 559 AAAGATATGGCAGAGAGAAAGCTTCAGTAGGAGAGGATACGGATCCGTTACAGGTATA 618
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
DB 619 TTTATGTATGCGGAAGATCTTGTGTACCCCGAAGATTACCGGATCGATGATCAGC 678
QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
DB 679 AACTTTCCCAAAAGAAATTAATGGAGATCAAAAGCGCAGATCATATGCAATGTTCTCC 738
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 739 AAGCCTCAAACTATGTCTCTTCTTGGAGATTGCAATATAAT 786

RESULT 12
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217

Alignment Scores:
Pred. No.: 1.56e-77 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x US-09-938-842A-217 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrPsrTyrTyrLeuLys 24
DB 22 CAACACGCTGCTTCTAGTACATGCTGCTGCCAGCGGTAGATCTAGCTGCTCGGTATA 81
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
DB 82 CCGAGCTCGAGGCTTCTGGCCACCGGTAAACCGCGGTAGATCTAGCTGCTCGGTATA 141
QY 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
DB 142 GACATGACCGAGTCAATCACAGATATATCCACATCGCAACAATATCTCAGAGCCATTGATG 201
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
DB 202 CAGCTAATGACCTCATTACACGATGATGATGATGATGATGATGATGATGATGATGATG 261
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
DB 262 GGTTCAGTTAGCTATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 321
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
DB 322 GTGACTGTATGTCGACAGACACCAACACTCCATCCTTCGTATGGATAAGCTAAGA 381
QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
DB 382 AAAGAAACTTCAGAGAGAAATGGTGTACACCGGTGTACAGC---GAGAAACCTGAT 438
QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
DB 439 TTTCCTAGCGAGTTTGGATTCTTGGACCAAGATTCATGGCCAAAGACTTGTATCAGTTG 498
QY 164 CysSerProGluAspLeuAlaSerLeuValArgProSerSerLeuPheMet 183
DB 499 TCTCCAGTCCAAAGATCTTGAATTTGGCGAAATTTGGTGGAGGCGCAACCATTTGATTAG 558
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
DB 559 AAAGATATGGCAGAGAGAAAGCTTCAGTAGGAGAGGATACGGATCCGTTACAGGTATA 618
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
DB 619 TTTATGTATGCGGAAGATCTTGTGTACCCCGAAGATTACCGGATCGATGATCAGC 678
QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
DB 679 AACTTTCCCAAAAGAAATTAATGGAGATCAAAAGCGCAGATCATATGCAATGTTCTCC 738
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 739 AAGCCTCAAACTATGTCTCTTCTTGGAGATTGCAATATAAT 786
```





```
QY      84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
      |||:::
      259 GGAATAACCTGCTCTCGCTCTGACATAATCCCTAGTAGAATGCTACTATTGTTTTTC 318
QY      104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123
      |||:::
      319 TTGACAGCTTTATGCCGACACAGAAACCTACTCTTATGTTTACAAAGCTAATC 378
QY      124 GluArgThrProAlaGlnPheLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
      |||:::
      379 AGAAGCGTCCACAGAGGATGGTTGGACACCGTGTGGAACTATGGGAACATGAA 438
QY      144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
      |||:::
      439 TGTCTCTAGAGTTTGCTCTTTTGGACCAAGTTTCATGGCCCAAGAAATTTGTATCAACTC 498
QY      164 CysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
      |||:::
      499 TCTCGGTCCAAGATCTTGAATTGGCAAAATGTTGGTGAGAGTAACCCCATCATTTACA 558
QY      184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
      |||:::
      559 AATAATCTGGCAGGCAAGAGCTTAGTGAGGAAGGTACGGTTACCGTTACACGTATA 618
QY      204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
      |||:::
      619 TATATTGTATGTGGAGGACATGGCGTACCCTGAGGATTCACAGTGGTGGATGATCAAG 678
QY      224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
      |||:::
      679 AACTTTCGCCCAAAAGAAATAATGAGATCAAAATGTGCAGATCATATGGCAATGTTCTCC 738
QY      244 GluProGlnLysLeuCysAlaSerLeuGluIleAlaHisLysTyr 259
      |||:::
      739 AAGCCTCACAAATATGTGCTCTCTCGTGGAGATGTCATGTAATAAT 786
```

## RESULT 15

```
US-10-437-963-62048
; Sequence 62048, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 62048
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63420C.1
US-10-437-963-62048
```

## Alignment Scores:

Pred. No.:	1,07e-74	Length:	1101
Score:	669.50	Matches:	134
Percent Similarity:	67.45%	Conservatives:	38
Best Local Similarity:	52.55%	Mismatches:	82
Query Match:	48.94%	Indels:	1
DB:	19	Gaps:	1

US-10-780-002-2 (1-260) x US-10-437-963-62048 (1-1101)

```
QY      5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrpTyrLysLeuLys 24
      |||:::
      81 AAGCACCTCGCTCTGCTGGTCCAGCGCGGTGCATCGCGGCTGGACCTATTTCAAGGTGGCG 140
QY      25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
      |||:::
      141 ACGGGCTTCGGTCCGGTCCGGGTACCGGTACCGGCGCGGACCTCGGCGCGCTCGGGCGTC 200
QY      45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
      |||:::
      201 GACCGCGCGCGTGGCGGAGGTCCGACGTTCGCGACTACACCGCGCGCTGCTGTGGCG 260
QY      65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
      |||:::
      261 CTCTCGGCTCCCTCCCTCCAGGGGAGAGGTGCTCTCGTCGGCCACAGGCTCGGGCGCG 320
QY      85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
      |||:::
      321 ATCAACGTCCGGCTGGCGCGGAGCTGTTCCCGGCAAGATCGCGCGCGCTGTTCTCTC 380
QY      105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
      |||:::
      381 TCGGCTTCATGTCGCGACACACGTCGCGCGCGTCCGACGTGCTCGAAAGTTCATCCAG 440
QY      125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
      |||:::
      441 GGGAAAGTGGCTGGGAC- --TGGATGGACACGAGATTTAAGCCCGCAGGATGCGAGGGCAAG 497
QY      145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
      |||:::
      498 CTCCCACTTTCATGTTGTTTCGGCGCGCAGATCGCCCAAGAAAGGCTGATCGAGCTGTC 557
QY      165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
      |||:::
      558 TCGCCAGAGGACGTCACCTCTCGCGGATCTCTGTAGGGTGAGCTCGATGTTCTGTGGAG 617
QY      185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
      |||:::
      618 GACCTGCAAAAGCAGCAGCCGTTCAACGAGGGCGCTACCGGCTCGGTCCGGAAGGTGTAC 677
QY      205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
      |||:::
      678 GTGGTCTAAACACGAGCTCGCCATACCGGAGGGGTTCCAGCGGTGATGATCGGGAAC 737
QY      225 IleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCysGlu 244
      |||:::
      738 AGCCCGGTGGATGAGGTGAAGGAGATCGATCGCGCGGACACCTCGTGATGCTCTCGAGG 797
QY      245 ProGlnLysLeuCysAlaSerLeuLeuIleAlaHisLysTyr 259
      |||:::
      798 CCCGACGAGTGGCGCGGTGCTCGCGACATCGCCGCAAGAGCTAC 842
```

Search completed: September 13, 2005, 00:37:19

Job time : 616 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 19:33:09 ; Search time 3112 Seconds  
(without alignments)  
3180.177 Million cell updates/sec

Title: US-10-780-002-2

Perfect score: 1368

Sequence: 1 MKEGHFVLVHGACHGWSW.....MLCEPQKLCASLEIAHKYN 260

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10780002/runat\_11092005\_104747\_26610/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002 @CGN\_1\_1\_5180 @runat\_11092005\_104747\_26610 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.\*  
1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hic:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gse1:.\*  
9: gb\_gse2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	79.0	922	CK269872	CK269872 EST115950
2	1045.5	76.4	886	CK260266	CK260266 EST115950
3	1045.5	76.4	907	CK266902	CK266902 EST112980
4	1045.5	76.4	914	CK270870	CK270870 EST116948
5	1045.5	76.4	966	CK260265	CK260265 EST106343
6	1000.5	73.1	781	BGI126704	BGI126704 EST472350
7	996.5	72.8	749	BQ514677	BQ514677 EST622092
8	978.5	71.5	646	AW621893	AW621893 EST312691
9	977.5	71.5	794	BM407979	BM407979 EST582306

10	965.5	70.6	784	4	BI434774	BI434774 EST337535
11	952.5	69.6	742	7	CK266903	CK266903 EST712981
12	867	63.4	957	7	CV267710	CV267710 WS02032.B
13	834	61.0	639	5	BQ111509	BQ111509 EST597085
14	819	59.9	796	7	CF808374	CF808374 P8H8034XM
15	798	58.3	865	7	CN190989	CN190989 UCRCS06.0
16	797	58.3	718	5	BQ853353	BQ853353 QGB20812
17	795	58.1	797	6	CB290370	CB290370 UCRCS01.0
18	795	58.1	821	6	CB290604	CB290604 UCRCS01.0
19	792	57.9	804	6	CB293837	CB293837 UCRCS01.0
20	787	57.5	781	7	CN190896	CN190896 UCRCS06.0
21	786	57.5	712	5	BQ852405	BQ852405 QGB17024
22	783	57.2	811	6	CB292675	CB292675 UCRCS01.0
23	780	57.0	850	6	CB292674	CB292674 UCRCS01.0
24	775	56.7	590	4	BI928709	BI928709 EST548598
25	773	56.5	834	7	CN191164	CN191164 UCRCS06.0
26	764	55.8	747	7	CN189384	CN189384 UCRCS06.0
27	762	55.7	797	7	CF837566	CF837566 UCRCS03.0
28	760	55.6	751	7	CN191146	CN191146 UCRCS06.0
29	759	55.5	535	2	BA433359	BA433359 EST399888
30	753	55.0	835	7	CF830288	CF830288 UCRCS01.0
31	750	54.8	740	7	CN191507	CN191507 UCRCS06.0
32	736	53.8	726	7	CF831499	CF831499 UCRCS01.0
33	733	53.6	733	7	CK110376	CK110376 N056C04.P
34	728	53.2	726	7	CF831872	CF831872 UCRCS01.0
35	716	52.3	623	5	BQ854991	BQ854991 QGB24M05
36	711	52.0	828	7	CF838704	CF838704 UCRCS03.0
37	708	51.8	701	7	CN185421	CN185421 UCRCS05.0
38	707	51.7	494	2	B8924354	B8924354 EST428123
39	703	51.4	747	1	AJ805825	AJ805825 AJ805825
40	701	51.2	476	1	AI772823	AI772823 EST253923
41	699	51.1	971	5	BQ481510	BQ481510 PV GBA003
42	697	51.0	719	7	CN190831	CN190831 UCRCS06.0
43	690	50.4	691	7	CN192311	CN192311 UCRCS06.0
44	690	50.4	911	3	CMS0A981	BX821370 Arabidops
45	689	50.4	525	1	AI775631	AI775631 EST256731

## ALIGNMENTS

CK269872 922 bp mRNA linear EST 03-AUG-2004  
EST115950 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POACP55 5' end, mRNA sequence.

CK269872  
CK269872.1 GI:39826850

EST.  
Solanum tuberosum (potato)

ORGANISM  
Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE  
1 (bases 1 to 922)  
Buell, C.R., Hart, A., Ziemann, V., Karamycheva, S.A. and Baker, B.

AUTHORS  
Generation of ESTs from abiotic stressed potato tissue

TITLE  
Unpublished (2003)

JOURNAL  
Other ESTs: EST115951

COMMENT  
Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/.

Seq primer: ART TAG GTG ACA CTA TAG.

Location/Qualifiers

1. 922

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POACP55"

/tissue\_type="abiotic stress treated leaf and root tissue"

```

/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cdna library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cdna library. RNA sample."
```

## ORIGIN

## Alignment Scores:

```

Pred. No.:      6,07e-117      Length:      922
Score:          1081.00      Matches:      207
Percent Similarity: 84.50%      Conservative: 22
Best Local Similarity: 76.38%      Mismatches: 30
Query Match:      79.02%      Indels:      12
DB:              7          Gaps:      2
```

US-10-780-002-2 (1-260) x CK269872 (1-922)

```

Qy      2  LysGluGlyLysHis-----PheValLeuVal 10
Db      4  AAAGAAAGAAAGCAAGCCCAAAATGAGGTTATGAGAAACATTTATGTTTGGTA 63

Qy      11 HisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLysPheProLeuLeuAlaAla 30
Db      64 CATGTGTCATGCCCATGGAAGTTGGTGGTATAGCTAAAGCCATTGTTAGAGGCTGCC 123

Qy      31 GlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThrAspLeuArgLysIleGlu 50
Db      124 GACCACACGGTCACCTGCCCTTGACATGCCGCTCTGGCATTGATTTGAGAAAAATAGAG 183

Qy      51 GluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGluLeuMetGluSerLeuSer 70
Db      184 GAACTTCGCACACTTGTGATTATACCGTGCCATTGATGGAGTTTATGGAAATCCCTTCCA 243

Qy      71 AlaAspGluLysValIleLeuValGlyHisSerLeuGlyGlyMetAsnLeuGlyLeuAla 90
Db      244 CAAGAAGAGAAAGGTATACGTAGGGGCATAGTTATGGTGGTATGAAATTTGGGACTTGCT 303

Qy      91 MetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAsp 110
Db      304 ATGGAAATAATACCCAAAAGATCTTGCTGCTGTTCTTCTGGCTGCTATTCTGCTGAT 363

Qy      111 SerValHisAsnSerPheValLeuGluGlnTyrAsnGluUArgThrProAlaGluAsn 130
Db      364 TCTGCTCATGCTCTCTATGTTTGTATCAGTACTTTTGAGAGGACACCAAAAGAGAAAT 423

Qy      131 TrpLeuAspThrGlnPheLeuProTyrGlySerProGluGluProLeuThrSerMetPhe 150
Db      424 TGGCTTGACACCAAAATTTGATCATATGTTTCCCTGGAAGAGCCCTCTGACATCCATGTTT 483

Qy      151 PheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAla 170
Db      484 TTTGGGCCCCAAGTTCTTGGCTCAACAACTTTATCATGATTATGCTCTCTCTCGAGGATTTGCA 543
```

```

Qy      171 LeuAlaSerSerLeuValArgProSerSerLeuPheMetGluAspLeuSerLysAlaLys 190
Db      544 TTAGCATCGTCATTGGTGAGCAACACCTCTCTGTTTATGGAAGATCTATCAAGGTGAAG 603

Qy      191 TyrPheThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAsp 210
Db      604 TATTTACGACGACGAGGTATGATCAGTCAGCAAGATTTATATATCTCAGCTGAGGAT 663

Qy      211 LysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsnIleGly----ValThrGlu 229
Db      664 AAAGTAACACCAAAAGAAATCCAGCAATGCGAAATTCACAACAATGGACCGTCATCGAA 723

Qy      230 AlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCysGluProGlnLysLeuCys 249
Db      724 GCAAGAGAGATCAAGGTGCTGATCATGTGTAATGCTATGTATGTCGCCCAACAACATTTGT 783

Qy      250 AlaSerLeuLeuGluIleAlaHisLysTyrAsn 260
Db      784 GACACTCTCTGGAGATTGCCATAATACAAT 816

RESULT 2
CK260266
LOCUS
DEFINITION
EST706344 potato abiotic stress cdna library Solanum tuberosum cdna
Clone POAB307 5' end, mRNA sequence.
ACCESSION
CK260266
VERSION
CK260266.1 GI:39817244
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 886)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other_ESTs: EST706342 EST706343 EST706345
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seg primer: CAG GAA ACA GCT ARG ACC.
FEATURES
Location/Qualifiers
1..886
/oranism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB307"
/tissue_type="abiotic stress treated leaf and root tissue"
/clone_lib="potato abiotic stress cdna library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
```



```

QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 162 GATTTGAGAAAAATAGAGAACTTTGCACACTTGTGTGATTATACCGTGCCATTGATGGAG 221
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 222 TTTATGGAAATGCTTCCACAGAGAGAGAGGTCATACCTAGTGGGCGCATGTATGGTGGT 281
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 282 ATGAATTTGGCACTTGCTATGAAAAATACCCAAAAAGATCTTTGCTGCTTTACTTG 341
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db 342 ACTGCTATTCTGCTGATCTACTACATGTCCTCTCTATGTTTGGATAAGTACTTGGAG 401
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 402 ACGACAACAAAGAGAGATTGGCTTGACACCCCAATTTGTATCATATGGTACCCCTGAAGAG 461
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 462 CTCGGACATCCATGCTTTTGGGCCCAAGTCTTGGCTCAAAAACCTTTATCAGTTATGC 521
QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db 522 CCTCCTGAGGATCTTTCATTAGCATCGTGTGGTGAGACCAACCTCTCTGTTTTCGAA 581
QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
Db 582 GATCTATCAAGGTGAAGTATTTCACGGACGAAGGGTTTGGATCATGTGAGAAAGATTTAT 641
QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
Db 642 ATATCTGCACCTGAGGATTAAGTATCATCCAAAAGATTCACGATGGCAATTCACAAAC 701
QY 225 IleGly---ValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 702 AATGGAACCGTCATCGAAGCAAGGAGATCAAAAGGTACTGTATCATATGCAATGCTATGT 761
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260
Db 762 AAGCCCAACGACTTGTGCGCACTCTCTTGGAGATGGCCCAATAATACAAT 812

```

## RESULT 4

```

CK270870
LOCUS CK270870 914 bp mRNA linear EST 03-AUG-2004
DEFINITION EST716948 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POACV79 5' end, mRNA sequence.

```

## ACCESSION

CK270870

## VERSION

CK270870.1

## KEYWORDS

EST

## SOURCE

Solanum tuberosum (potato)

## ORGANISM

Solanum tuberosum

## REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum.

## AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

## TITLE

Generation of ESTs from abiotic stressed potato tissue

## JOURNAL

Unpublished (2003)

## COMMENT

Other ESTs: EST716949

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@igr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

Location/Qualifiers

1..914

/organism="Solanum tuberosum"

```

/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clones="POACV79"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/notes="vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
12 hr and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 9.59e-113 Length: 914
Score: 1045.50 Matches: 197
Percent Similarity: 85.99% Conservative: 24
Best Local Similarity: 76.65% Mismatches: 35
Query Match: 76.43% Indels: 1
DB: 7 Gaps: 1
US-10-780-002-2 (1-260) x CK270870 (1-914)

```

```

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
Db 96 AAACACCTTTGTTGGTACATGGTGTCATGCCATGGAGTTGGTGTGTGTAAGCTAAG 155
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 156 CCATTTGTAGAGGCTGCGGCCACACGGTCACTGCCCTTGACATGGCGCCCTCTGGCATT 215
QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 216 GATTGAGAAAAATAGAGAACTTTGCACACTTGTGTGATTATACCGTCCCATTTGATGGAG 275
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 276 TTTATGGAAATGCTTCCACAGAGAGAGAGGTCATACCTAGTGGGCGCATGTATGGTGGT 335
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 336 ATGAATTTGGCACTTGCTATGAAAAATACCCAAAAAGATCTTTGTTGCTGTTTACTTG 395
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db 396 ACTGCTATTCTGCTGATCTACTACATGTCCTCTCTATGTTTGGATAAGTACTTGGAG 455
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 456 ACGACAACAAAGAGGATTTGGCTTGACACCCCAATTTGTATCATATGGTACCCCTGGAAGAG 515
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 516 CCTCGGACATCCATGCTTTTGGGCCCAAGTCTTGGCTCAAAAACCTTTATCAGTTATGC 575

```



DEFINITION EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
 cTOP13K13 5' sequence, mRNA sequence.  
 ACCESSION BG126704  
 VERSION BG126704.1 GI:12626892  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 781)  
 AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,  
 Hansen, C., Romin, C. and Tanksley, S.  
 TITLE Generation of ESTs from tomato shoot/meristem tissue  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: CUGI  
 Clemson University  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>.  
 Location/Qualifiers  
 1..781  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOP13K13"  
 /tissue\_type="shoot/meristem"  
 /dev\_stage="developing shoots from 4-6wks old plants"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato shoot/meristem"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Small expanding leaves from the growing tip were  
 taken from greenhouse plants (4-6wks old TA496). Tissue  
 was immediately frozen in liquid nitrogen."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.62e-107 Length: 781  
 Score: 1000.50 Matches: 188  
 Percent Similarity: 87.1% Conservative: 23  
 Best Local Similarity: 77.6% Mismatches: 30  
 Query Match: 73.1% Indels: 1  
 DB: 4 Gaps: 1  
 US-10-780-002-2 (1-260) x BG126704 (1-781)  
 QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTTPSerTTPYr 21  
 DB 48 AAACAGGAGGACACTTTGTTGGTACATGTGTGATGCCATGGAGGTGGTGGTGGTAC 107  
 QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41  
 DB 108 AACTAAAGCCCTTCTAGAGGTTGCAGGCCACCAAGGTCACAACTCTTGATTGGCAGCC 167  
 QY 42 SerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
 DB 168 TCTGGTATTGATTGAGAAAAATAGACCAACTTCACACATTTTCATGATTACACTTTGGCA 227  
 QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81  
 DB 228 TTGATGGAATGATGGAATCTCTTCCACAGAGGAGAAAGTCATCTAGTTGGACATAGT 287  
 QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101  
 DB 288 CTTGGTGGTATGAATTTGGACTTGTATGGAAAAATACCCACAAAAGATCTATCTTGT 347  
 QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121  
 DB 348 GTTTCTTGGCGCTTCATGCTGATCTATTCTATCTGCTCTCTGCTCTGATGATCAG 407  
 QY 122 TyrAsnGluArgThrProAlaGluAsnTTPLeuAspThrGlnPheLeuProTyrGlySer 141

Db 408 TATTTTGAGAGGTGCAACCAATGATGTTGGTGGACACCAATTTGTATCATATGTTGCC 467  
 QY 142 ProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyr 161  
 Db 468 CATGAAGAGCCTGTGCCATCATATTTTGGGCCCAAGTTCTGGCTTACAATCTTTAT 527  
 QY 162 GlnLeuCysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeu 181  
 Db 528 CAATATGCGCTCTCTGAGGATGTTGCAATGATATCATATGGGAGAGCAAGTCTTTA 587  
 QY 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201  
 Db 588 TTCTAGAGATTGTTCAAAATCAAGTATTGTGACAGATGAAGGATATGGATCAGTGAAG 647  
 QY 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTTPGln 221  
 Db 648 AAGTTTATATAGTATGCACAGATGATAAATCTCTACCAAAAGAGTTTCAAAAATGGCA 707  
 QY 222 IleAspAsnIle---GlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240  
 Db 708 ATTGACCAACATTAAATGATCATGATAAACAAGAAATTGAAGTCTGATCATATGGCA 767  
 QY 241 MetLeu 242  
 Db 768 ATGCTA 773  
 RESULT 7  
 BQ514677  
 LOCUS BQ514677  
 DEFINITION EST622092 Generation of a set of potato cDNA clones for microarray  
 analyses mixed potato tissues Solanum tuberosum cDNA clone STMIN36  
 3' end, mRNA sequence.  
 ACCESSION BQ514677  
 VERSION BQ514677.1 GI:21373546  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 749)  
 AUTHORS Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,  
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and  
 Karamycheva, S.A.  
 TITLE Generation of a set of potato cDNA clones for microarray analyses  
 JOURNAL Unpublished (2002)  
 COMMENT Other ESTs: EST622091  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: [potato-array@tigr.org](mailto:potato-array@tigr.org)  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>  
 Seq primer: T7.  
 FEATURES  
 Location/Qualifiers  
 1..749  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec or Binjite"  
 /db\_xref="taxon:4113"  
 /clone="STMIN36"  
 /tissue\_type="mixed tissues"  
 /lab\_host="SOLR"  
 /clone\_lib="Generation of a set of potato cDNA clones for  
 microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Combination of untreated and Phytophthora  
 infestans-treated libraries of stolons, leaves, leaflets,  
 axillary buds of stem explants, petioles, germinating  
 eyes, tubers, or roots."  
 ORIGIN





```

Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 492 TACCAATATTGCTCTCCCTGA----- 511
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 512 -----GATTTGCAAGGAAGAAAGTACTTGGAGTGAAGTATATGATCGGTG 559
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr 220
Db 560 AAGAGAGTTTACATAGTATGCACTGAGGACAAAGGCATACCAACAGAAATTTCAACAATGG 619
Qy 221 GlnIleAspAsnIleGlyValThrGlu 229
Db 620 CAAATTGACAACATTGGTGTGACAGAA 646

RESULT 9
LOCUS BM407979 794 bp mRNA linear EST 10-MAR-2003
DEFINITION EST582306 potato roots Solanum tuberosum cDNA clone cPRO32J8 5'
end, mRNA sequence.
ACCESSION BM407979
VERSION BM407979.1 GI:18259609
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Teai,J., Van Aken,S.,
Utterback,T., Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C.,
Tanksley,S. and Baker,B.
Generation of ESTs from potato roots
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3
FEATURES
Location/Qualifiers
1..794
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPRO32J8"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/clone_lib="potato roots"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."
ORIGIN
Alignment Scores: 8.81e-105 Length: 794
Pred. No.: 977.50 Matches: 185
Score: 85.54% Conservative: 22
Percent Similarity: 76.45% Mismatches: 34
Best Local Similarity: 71.45% Indels: 1
Query Match: 4 Gaps: 1
DB:

US-10-780-002-2 (1-260) x BM407979 (1-794)
Qy 5 LysHisPheValLeuHisGlyAlaCysHisGlyTyrTrpSerTrpTyrLysLeuLys 24

```

```

Db 67 AACACATTTGTTGGTACATGTCATGGAGAGTTGGTGTGCTAAGCTAAAG 126
Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 127 CCATTGTTAGAGGTCGCCGCCACACGGTCCTCAGTCCTTCACATGCGCGCTCTGGGCATT 186
Qy 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 187 GATTTGAAAAAATAGAGGAACCTTTGCACACTTGTGTTGATTATACCGTGCCATTGATGGAG 246
Qy 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 247 TTTATGGAATGCCTTCACCAAGAGAGAGTCTACTAGTGGGCGCATAGTTATGTTGGT 306
Qy 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 307 ATGAATTTGGCACCTTGCTATGGAATAATACCAAAAAAGATCTTTGTGCTGTTTACTTG 366
Qy 105 AlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsnGlu 124
Db 367 ACTGCTATTTCGCTCTGATTCTACTCACATGTCCTCTCTATGTTTGGATAAGTACTTGGAG 426
Qy 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 427 ACGACAAACAAAGAGGATTGCTTGACACCAATTTGTATCATATGTTACCCCTGAAGAG 486
Qy 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 487 CCTCGGCATCATGCTTTTGGGCCCAAGTCTTGGCTCAAAAACTTTATCACTGTTATGC 546
Qy 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db 547 CCTCCTGAGGATCTTTCATTAGCATGTCGTGTTGGTGAGCAACCTCTCTGTTTATCGAA 606
Qy 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
Db 607 GATCATCAAGGTGAAGTATTTCCAGCAGCAAGGGTTTGGATCAGTCAAGAAAAGTTTAT 666
Qy 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyrGlnIleAspAsn 224
Db 667 ATAATCTGCACTGAGGATAAAGTATCAACCAAAAGATTCACGAATGCAAAATGACCCAC 726
Qy 225 IleGly---ValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 727 CATGGGACCGTCATCGAAGCAAGAGATCAAGGTACTGATCATATGGCAATGCTATGT 786
Qy 244 GluPro 245
Db 787 AAGCCC 792

RESULT 10
LOCUS BI434774 784 bp mRNA linear EST 30-APR-2003
DEFINITION EST537535 P. infestans-challenged potato leaf, compatible reaction
Solanum tuberosum cDNA clone PPCBU48 5' sequence, mRNA sequence.
ACCESSION BI434774
VERSION BI434774.1 GI:15259464
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans. Compatible Interaction
Unpublished (2000)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA

```





```
Db 159 AAGAAATCAGAGCATTTTGTCTAGTACATGGAGCTTGCATGGCGCTTGGTGTGGCAA 218
Qy 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
Db 219 AAGTTCAAAACGCTGCTGTAGTACAGCAAGTAACAGGCTCACGGTGTCTGACCTTGTGCT 278
Qy 42 SerGlyThrAspLeuArgLysIleGluLeuLeuArgThrLeuTyrAspTyrThrLeuPro 61
Db 279 TCAGGGCCCAACATGAAGCCATCCAAGATGTAGAACAACACTTGATCAATATACGGAGCCT 338
Qy 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Db 339 TTCTTAGAGTTTCTGGCTCATTAACAACCAAGAGAGAGTCACTTCTAGTAGGCACAGC 398
Qy 82 LeuGlyGlyMetAsnLeuLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Db 399 CTAGGAGGTTTGGATTGGCTCTGTATGGAAGTTCCAGAGAGAGATTGCTGTGCT 458
Qy 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
Db 459 GTTTCCTTATCAGCTTTCATGCCAGATACACACAAAGCCATCATTTGCTTGGATCAG 518
Qy 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
Db 519 TATAACGAGAGGACCCCGCGGATTCCTGTTGGACACTCAATTTTACCATACAGCACT 578
Qy 142 ProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyr 161
Db 579 TCTCAAAAGTCATCTCAACAACATGTCTTTTGGACCCAAATTTCTATCTCCCAAGCTCTAT 638
Qy 162 GlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerLeu 181
Db 639 CAGCTAAGCCCACTCAGGATCTTGGACGAAAGACATATGTATGAAGCAGATCACCTG 698
Qy 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201
Db 699 TTTCTGTATGATTGTCAAGGCAACAGATTTCTCCACGACGGCTATGGTCAGTCAAA 758
Qy 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGln 221
Db 759 CGAGTGTATGTTATCTCGCATGAAGATTTAGCGATACACAGAGAGTTTCAACGCTGGATG 818
Qy 222 IleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMet 241
Db 819 ATTGAANAAGCTCTGTTGAGAAGTTATGGAATTTGAAGGTGCAGACCATATGTTATG 878
Qy 242 LeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 879 TTCAGCAAGCCACAAAGACTCTTCCATTGTCTCTCAGAGATAGCAAAATAACAT 932
```

## RESULT 13

```
BQ111509
LOCUS EST597085 mixed potato tissues Solanum tuberosum cDNA clone STMCB21
DEFINITION 5', end, mRNA sequence.
```

```
ACCESSION BQ111509
VERSION BQ111509.2 GI:21914979
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 639)  
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karanicheva, S.A.

Generation of a set of potato cDNA clones for microarray analyses  
Unpublished (2002)  
On Apr 17, 2002 this sequence version replaced gi:20163471.  
Contact: Robin Buell  
The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igr.org  
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.

## FEATURES

Location/Qualifiers  
source 1..639  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec or Binjete"  
/db\_xref="taxon:4113"  
/clone="STMCB21"  
/tissue\_type="mixed tissues"  
/lab\_host="SOLR"  
/clone\_lib="mixed potato tissues"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."

## ORIGIN

Alignment Scores:  
Pred. No.: 6,32e-88 Length: 639  
Score: 834.00 Matches: 157  
Percent Similarity: 86.57% Conservative: 17  
Best Local Similarity: 78.11% Mismatches: 27  
Query Match: 60.96% Indels: 0  
DB: 5 Gaps: 0

US-10-780-002-2 (1-260) x BQ111509 (1-639)

```
Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTyrTyrLysLeuLys 24
Db 35 AAACACTTTGTTTGGTATCATGGTGCATGCCATGGAAGTTGGTGTGATTAAGTAAAG 94
Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 95 CCATTGTTAGAGCTGCGGCCACACGCGTCACTGCCCTTGACATGCCGCGCTCTGGCAT 154
Qy 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 155 GATTTGAGAAAATAGAGAACTTTTGGACACTTTGTTGATTATACCGTGCATTGATGGAG 214
Qy 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 215 TTTATGGAATGCCTTCCAAAGAGAGAGAGGTCACTACTAGTGGGCAATAGTTATGTTGT 274
Qy 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 275 ATGAATTTGGCACTTCTGTATGGAANAATACCAAAAAGATCTTTGTTGCTGTTACTTG 334
Qy 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db 335 ACTGCTATTCTGCCTGATTCTACTCACATGCTCTCTATGTTTGGATAAGTACTTGGAG 394
Qy 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 395 ACGACAACAAGAGAGATTGGCTTGACACCAATTTGTATCATATGTTACCCCTGAAGAG 454
Qy 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 455 CTGCGCATCCATCTTTTGGGCCCCAAGTTCTTGGCTCAAAAACCTTTATCAGTTATGTC 514
Qy 165 SerProGluAspLeuAlaAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db 515 CCTCTCTGAGGATCTTTTATTAGCATCGTGTGTTGGTGGAGACCAACCTCTCTGTTATCGAA 574
Qy 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
Db 575 GATCTATCAAGGTGAAGTATTTTACGAGACGAAAGGTTTGGATCATGATGAAGAAGTTTAT 634
```

```

QY      205 ile 205
Db      635 ATA 637

RESULT 14
CF808374
LOCUS   796 bp mRNA linear EST 15-APR-2004
DEFINITION psHB034xM24f USDA-IPAFS:Expression of Phytophthora sojae genes
           during infection and propagation Glycine max cDNA clone sHB034M24
           5, mRNA sequence.
ACCESSION CF808374
VERSION   CF808374.1 GI:37996785
KEYWORDS EST.
SOURCE   Glycine max (soybean)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
           Glycine.
REFERENCE 1 (bases 1 to 796)
AUTHORS   Tyler,B.
TITLE     Tyler,B. Not Published
JOURNAL   Unpublished (2003)
COMMENT   Contact: Tyler B
           Tyler lab
           VBI
           1880 Pratt Dr., Blacksburg, VA 24061, USA
           Tel: 540-231-7318
           Email: bmtylev@vt.edu
           PCR Primers
           FORWARD: BK reverse primer
           BACKWARD: BK reverse primer
           Plate: 034 row: M column: 24
           Seq primer: BK reverse primer
           High quality sequence stop: 796.
           Location/Qualifiers
           1..796
           /organism="Glycine max"
           /mol_type="mRNA"
           /cultivar="Harosoy"
           /db_xref="taxon:3847"
           /clone="sHB034M24"
           /tissue_type="Phytophthora sojae-infected hypocotyl"
           /cell_line="Phytophthora sojae culture P6497"
           /dev_stage="48 hr. post infection stage"
           /clone_lib="USDA-IPAFS:Expression of Phytophthora sojae
           genes during infection and propagation"
           /note="Vector: pBK-CMV; Site_1: EcoR1; Site_2: Xho1"

ORIGIN
Alignment Scores:
Pred. No.: 5,23e-86 Length: 796
Score: 819.00 Matches: 154
Percent Similarity: 76.56% Conservative: 42
Best Local Similarity: 60.16% Mismatches: 55
Query Match: 59.87% Indels: 5
DB: 7 Gaps: 1

US-10-780-002-2 (1-260) x CF808374 (1-796)

QY      5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTyrTyrIleuLys 24
Db      39 AAGCACTATGTTCTGGTGCATGGGCGATGCCATGGAGCTTGGTGTGATTAAGCTCAAG 98
QY      25 ProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db      99 CCACGCTTGATCTGCAGCGCCATAGGTACAGTACTTGACCTGTCAGCTTCGGAACC 158
QY      45 AspLeuArgGlyLeuGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db      159 AACATGAAGAAAATTGAAGATGTGTATCTTCTCAGAGTATCTCGCGCTTTGTTGCAG 218
QY      65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84

Db      219 CTAATGGCCACAAATTCCTCAATGAGAAGTAGTTCTAGTTGGTCACAGCCTTGGAGGG 278
QY      85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db      279 CTGAACATAGCACATTGCAATGGAGAAATTCACAGAAAGGTAGCAGTTGGTGTGTTCTTA 338
QY      105 AlalaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db      339 ACAGCTTTTGTCTCAGACACTGAACACACCCATCTTATGCTTTGGAAAGATACAATGAG 398
QY      125 ArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db      399 AGGACCCCGTTAGCTGCATGGTTAGACACTGAATTTGCTCCAGCGGAAACAAA----- 452
QY      145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db      453 -----ACATCAATGTTCTTCGGCCCCCAACTCTTGTCCGACAAGCTCTACCAACTGTCC 506
QY      165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db      507 CCTATTGAGGACTTGGAAATGGCCAAGACTTTAGCAAGGCCATCATCGCTCTTCATGGAA 566
QY      185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPhe-GlySerValLysArgValTy 204
Db      567 TATTGTTTGCACGAGGAGGACCTTGGAAATTCATTCGAAATATCAGCTCTTGTGATGATCCAAA 686
QY      204 rIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAs 224
Db      627 TATTGTTTGCACGAGGAGGACCTTGGAAATTCATTCGAAATATCAGCTCTTGTGATGATCCAAA 686
QY      224 nIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCysGlu 244
Db      687 TGTTGGGTTCAATGACGTTGTAGAGGTCANAGACGACATCATATGTTATGCTNTGCAA 746
QY      244 uProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db      747 GCCACAGAACTATTCGATTCCTCCAGCAGATAGCGACTAATAT 792

RESULT 15
LOCUS   CN190989
DEFINITION UCRCS06_0004G06_r Washington Navel Orange Stored Fruit Rind cDNA
           Library Citrus sinensis cDNA clone CS_WEC0004G06, mRNA sequence.
ACCESSION CN190989
VERSION   CN190989.1 GI:46215928
KEYWORDS EST.
SOURCE   Citrus sinensis
ORGANISM Citrus sinensis
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE 1 (bases 1 to 865)
AUTHORS   Close,T.J., Roose,M.L., Arpaia,M.L., Federici,C.F., Fenton,R.D.,
           Wanmaker,S., Focht,E., Sievert,J., Robinson,P., Kim,H.R.,
           Kudrna,D., Stum,D., Yost,D. and Wing,R.
           Development of EST Resources and New Genetic Markers for California
           Citrus - Washington Navel Orange Stored Fruit Rind
           Unpublished (2004)
JOURNAL   Contact: Timothy Close
COMMENT   Department of Botany & Plant Sciences, University of California
           Riverside, CA, 92521-0124
           Tel: 9097873318
           Fax: 9097874437
           Email: timothy.close@ucr.edu
           Seq primer: T3.
           Location/Qualifiers
           1..865
           /organism="Citrus sinensis"
           /mol_type="mRNA"
           /cultivar="Washington navel"
           /db_xref="taxon:2711"
           /clone="CS_WEC0004G06"

```

```

/tissue_type="Rind"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TJCl21"
/clone_lib="Washington Navel Orange Stored Fruit Rind cDNA
Library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Fruits were collected
January-March 2003 (Federici, Roese lab; Focht, Sievert &
Robinson, Arpaia lab). Four samples related to storage
conditions were produced: 1) fresh-picked in Mentone
(Arnott Brothers Enterprises, Mentone, CA), 2) after 21
days storage at 5C at Kearney then transported to UC
Riverside on ice, 3) after 5 additional days storage at
11C at Kearney, sampled immediately, 4) fruit grown in
southern CA were obtained from Redlands Foothill Packing
House after commercial packing, X-ray irradiated at 300 Gy
by Surebeam, then stored 1 day at ambient temperature.
Rind tissue (juice vesicles) were collected. Tissues were
snap frozen and then stored at -80C until further
processing. Fenton (Close lab) purified RNA by the phenol
method described in J. Japanese Soc. Hort. Sci. 1996. 64
(4): 809-814, purified poly(A) mRNA using a PolyAtrack
mRNA Isolation System IV (Promega), produced a primary
cDNA library using a lambda ZAP XR cDNA Synthesis Kit
(Stratagene), then mass-excised one million pfu from the
primary library to produce a phagemid population.
Phagemids were plated, plasmid DNA purified, cDNA clones
archived, and DNA sequences determined bi-directionally
using an ABI3730 at the Arizona Genomics Institute,
University of Arizona (Kim, Kudrna, Stum, Yost, Wing).
Chromatogram files were downloaded by FTP to UC Riverside
(by Close), then processed at UC Riverside (by Wanmaker,
Close lab) using the HarVest pipeline
(http://harvest.ucr.edu) to remove vector and cloning
oligo sequences and various contaminants, and to trim to a
high quality region. Sequences that retained a phred 17
region of at least 100 bases were deposited to GenBank."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1.82e-83 Length: 865
Score: 799.00 Matches: 152
Percent Similarity: 73.64% Conservative: 38
Best Local Similarity: 58.91% Mismatches: 66
Query Match: 58.33% Indels: 2
DB: 7 Gaps: 1

```

US-10-780-002-2 (1-260) x CN190989 (1-865)

```

QY 4 GlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrLysLeu 23
Db 3 GGGAGCATTTGTTCTAGTTCATGGAGTAACCATGGAGCATGGTGGTACAACTG 62

QY 24 LysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGly 43
Db 63 AAAGCAAGCTGGTGGCGGGGTACCGGGTGACGGCTGGACCTAGCCGCTCGGC 122

QY 44 ThrAspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 123 ATCAACATGAAGAGAATTGAGGATGTGCACACATTCATGCATGAGCCCTTGATG 182

QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 183 GAGGTTTGGCATCTCTCTCCGGAAGAAAGGTCTACTTGTGGACACAGTCTTGGG 242

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaValPhe 103
Db 243 GCGTCACCTTGGCCCTTGGCCGCGACAAATCCCAACAAATCTCCGTGGCTGTTTC 302

QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
Db 303 GTAACGTGCAATTCATGCTGCACACACACCGCCCATCTTTGTTTGGAGCAGTATTCT 362

```

```

QY 124 GluArgThrProAlaGlu-----AsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
Db 363 GAGAGATGGGAAAGAGAGACGACAGCTGGTTGGACACTCAATTTTCAAAATGTGACGG 422

QY 142 ProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyr 161
Db 423 TCAAAATCCATCTTCACATTTCCATGCTTTTCGGCGCGAGTCTTCTGACTATCAAGATCTAT 482

QY 162 GlnLeuCysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeu 181
Db 483 CAGCTTTGTCTCTCCCTGAGGATCTGGAGCTGGCCAAAGATGTTGGTGAGGCCAGATCAATG 542

QY 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201
Db 543 TTTATAGACAACCTTATCGAAGGAAAGTAAGTTTCAGCGATGAAGGATACGGATCTGTTAAG 602

QY 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGln 221
Db 603 CGAGTTTATCTTGTATGCGAAGAGGATATTTGGTCTCTCCCTAAGCAATTTTCAGCACTGGATG 662

QY 222 IleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMet 241
Db 663 ATCCAAAATATCTCCGTTTAATGAAGTGTGAGAGATCAAGGGCGGTGATCACAATGGCAATG 722

QY 242 LeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 723 CTTTCGACCCACACAGAACTTTGCGATGTTCTGTCTCAGATTTCTCTTAAAGTAT 776

```

Search completed: September 12, 2005, 22:47:38  
Job time : 3122 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 12, 2005, 17:23:43 ; Search time 482 Seconds  
(without alignments)  
3193.221 Million cell updates/sec

Title: US-10-780-002-2  
Perfect score: 1368  
Sequence: 1 MKEGKHFVLVHGAGCHGWSW.....MLCEPQKLCASLEIAHKYN 260

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO spool\_p/US10780002/runat\_11092005\_104746\_26588/app\_query.fasta\_1.455  
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002@cgn\_1\_1.708 @runat\_11092005\_104746\_26588 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	Description
1	1368	100.0	1079	8 ADA26498 Tobacco s
2	768	56.1	792	6 ABZ13515 Arabidops
3	768	56.1	816	3 AAC42344 Arabidops
4	768	56.1	918	8 ADA26530 Arabidops
5	696	50.9	870	3 AAC34411 Arabidops

6	695	50.8	792	6	ABZ13746 Arabidops
7	695	50.8	864	3	AAC47519 Arabidops
8	695	50.8	950	8	ADA26531 Arabidops
9	690	50.4	792	3	AAC43133 Arabidops
10	690	50.4	792	6	ABZ12412 Arabidops
11	690	50.4	792	12	ADN74620 Thale cre
12	690	50.4	1152	8	ADA26532 Arabidops
13	680.5	49.7	1109	8	ADA26533 Arabidops
14	665.5	48.6	1046	8	ADA26535 Arabidops
15	656	48.0	984	3	AAC49239 Arabidops
16	654.5	47.8	834	8	ADA26537 Arabidops
17	645	47.1	783	8	ADA26536 Arabidops
18	640.5	46.8	770	8	ADA67827 Arabidops
19	640.5	46.8	771	6	ABZ13473 Arabidops
20	640.5	46.8	823	3	AAC42529 Arabidops
21	640.5	46.8	928	8	ADA26538 Arabidops
22	635	46.4	909	3	AAC47800 Arabidops
23	635	46.4	984	8	ADA26534 Arabidops
24	635	46.4	1056	3	AAC47895 Arabidops
25	632.5	46.2	795	8	ADA70991 Rice gene
26	603.5	44.1	1325	3	AAC47916 Arabidops
27	590	43.1	828	3	AAC41733 Arabidops
28	590	43.1	908	8	ADA26539 Arabidops
29	589	43.1	903	3	AAC34141 Arabidops
30	585	42.8	817	2	AAT86824 Hevea bra
31	585	42.8	1091	2	AAT36351 Hevea bra
32	574.5	42.0	978	8	ADA69955 Rice gene
33	548	40.1	860	8	ADA71305 Rice gene
34	530.5	38.8	777	12	ADQ16305 Nucleotid
35	530.5	38.8	777	12	ADQ16306 Nucleotid
36	501.5	36.7	783	8	ADA70766 Rice gene
37	438	32.0	720	8	ADA70729 Rice gene
38	426.5	31.2	422	3	AAC43356 Arabidops
39	393.5	28.8	1550	8	ADA26540 Arabidops
40	385.5	28.2	1155	10	ABT23286 Seed deve
41	380	27.8	1815	3	AAC44497 Zea mays
42	377.5	27.6	1665	8	ADA26542 Arabidops
43	368	26.9	1268	3	AAC32766 Arabidops
44	368	26.9	1751	8	ADA26543 Arabidops
45	367	26.8	771	3	AAC43020 Arabidops

ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
FT CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX Klessig DF, Kumar D;  
 XX WPI; 2003-278578/27.  
 DR P-PSDB; ADA26499.  
 XX  
 PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.  
 XX  
 PS Claim 1; Fig 4; 98pp; English.  
 XX  
 CC The invention relates to a novel isolated 1079 bp salicylic acid-binding  
 CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening  
 CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the SABP2 protein from tobacco plants.  
 XX  
 SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,36e-149 Length: 1079  
 Score: 1368.00 Matches: 260  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-780-002-2 (1-260) x ADA26498 (1-1079)

QY 1 MetLysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrp 20  
 DB 42 ATGAAGGAGGAAACACATCTTGTATTAGTACATGTCATGTCATGAGGTTGGATTGG 101  
 QY 21 TyrLysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 40  
 DB 102 TACAAGCTAAAGCCACTGCTAGAAGCTGCAGGCCATAAGGTTACAGCCCTTGATTAGCA 161  
 QY 41 AlaSerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
 DB 162 GCTTCTGGCAGCTGATTGTGAGAAAATAGAGAGCTTCGCACACTTATATGATTTACTTTG 221  
 QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
 DB 222 CCATTGATGGAGTTGATGGAATCTCTTTCAGCAGATGAGAGGTTATATTAGTGGGCAT 281  
 QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 DB 282 AGCTTTGGTGGTATGAATTTGGGACTTGTATGAAAGATATCCACAAAAGATCTATGCT 341  
 QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
 DB 342 GCTGTTTTCTGGCTGCTTCATGCTGATCTCTTCACAACTCCTCTCTTTGTTTGGAA 401  
 QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 DB 402 CAGTATATGAGCGAGCGCCAGCCCGAGGATTTGGTACTACTCAGTTTACCATATGGT 461  
 QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
 DB 462 TCCCTTGAAGAGCCACTGATCCATCTGTTTTTGGCCCAAGTCTCTTGCTCACAAGCTC 521  
 QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180

DB 522 TACCAGCTATGCTCTCCTGAGGATCTTGTAGCATCATCTGGTGAGACCAAGCTCT 581  
 QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 DB 582 TTGTTTATGGAAGACCTATCGAAGCCCAAGTAATTTCAAGATGAACGGTTTGGATCAGTG 641  
 QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrp 220  
 DB 642 AAGAGAGTTTACATTTGTGTGCTGAGGATTAAGGCATACCAAGAAATTCACGCGATGG 701  
 QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
 DB 702 CAAATTGACACATTTGGTGTCTCACTGAAGCAATAGAGATTAAGGTGCTGATCACATGCA 761  
 QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrTrp 260  
 DB 762 ATGCTATGCGAGCCCAAAAACCTTTGGCGCTCTCTCTTGGAAATTTGCCCATATAACAAC 821

RESULT 2  
 ABZ13515  
 ID ABZ13515 standard; DNA; 792 BP.  
 XX  
 AC ABZ13515;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1320.  
 XX  
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200216655-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 24-AUG-2001; 2001WO-US026685.  
 XX  
 PR 24-AUG-2000; 2000US-0227866P.  
 PR 26-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Harper JF, Krepes J, Wang X, Zhu T;  
 XX  
 WIPI; 2002-304127/34.  
 XX  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX  
 PS Claim 144; SEQ ID NO 1320; 577pp + Sequence Listing; English.  
 XX  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 SQ Sequence 792 BP; 217 A; 167 C; 196 G; 212 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.81e-79 Length: 792  
 Score: 768.00 Matches: 149  
 Percent Similarity: 71.04% Conservative: 35

Best Local Similarity: 57.53%		Mismatches: 73	
Query Match: 56.14%		Indels: 2	
DB: 6		Gaps: 2	
US-10-780-002-2 (1-260) x AB213515 (1-792)			
QY	2	LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyr	21
DB	13	AAGAGGAACAACATTTGTACTAGTACATGGTTCGTGCCATGGCGGTGGCTGGTAC	72
QY	22	LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla	41
DB	73	AAGTTAAGCGCTGTAGAGCGGTGGGCCACCGCGTAACTGCTGTGGACTTAGTGCC	132
QY	42	SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu	60
DB	133	TCCGGAATAGACAACAGAGGTGCATCTGCATATCCCCACATCGCAACAATCTCGGAG	192
QY	61	ProLeuMetGluLeuMetGluSerLeuSerAlaAspGlyLysValIleLeuValGlyHis	80
DB	193	CCATTGACGAGCTCTGACCTCATTTGCCAATGATGAAGGTTGTGCTCGTTGGTCTAC	252
QY	81	SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla	100
DB	253	AGCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAGTTTCCCGAAAAATCTCTGTC	312
QY	101	AlaValPheLeuAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu	120
DB	313	GCTGTATTCTTGACTCTTTCATGCGGACACCGAACACTCCATCCTTCGTTGGAC	372
QY	121	GlnTyrAsnGluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGly	140
DB	373	AGTTTGGAGCAACATGCTCAAGAACATGATGGGACCGAATTCGAACCTTATGTT	432
QY	141	SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu	160
DB	433	TCAGACAATTCGGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC	489
QY	161	TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer	180
DB	490	TACCAGCTTTCCTCCAGTCTGAGGATCTTGAACCTGGGATTACTTTTAATGAGCCAGGATCG	549
QY	181	LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal	200
DB	550	TTATTATTAAACGATTATCGAAGATGAAGAAACTTCTCGGATGAAGGATATGGTCTGTT	609
QY	201	LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr	220
DB	610	CCTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTTCCAGAAAGAACGCCAGAGATCG	669
QY	221	GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla	240
DB	670	ATGATTGATAAATTTCCGGTGAATTTAGTATGATGGAGATGGAGACAGATCATATGCCA	729
QY	241	MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr	259
DB	730	ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAAATCGCGGACAAATTC	786

RESULT 3  
AAC42344  
ID AAC42344 standard; DNA; 816 BP.  
XX AAC42344;  
AC AAC42344;  
XX  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS

XX	EP1033405-A2.
PN	
XX	06-SEP-2000.
PD	
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	28-APR-1999; 99US-0130891P.
PR	30-APR-1999; 99US-0131445P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.
PR	18-JUN-1999; 99US-0139460P.
PR	18-JUN-1999; 99US-0139461P.
PR	18-JUN-1999; 99US-0139462P.
PR	18-JUN-1999; 99US-0139463P.
PR	18-JUN-1999; 99US-0139750P.
PR	18-JUN-1999; 99US-0139763P.
PR	21-JUN-1999; 99US-0139817P.
PR	22-JUN-1999; 99US-0139899P.
PR	23-JUN-1999; 99US-0140353P.
PR	23-JUN-1999; 99US-0140354P.
PR	24-JUN-1999; 99US-0140695P.
PR	28-JUN-1999; 99US-0140823P.
PR	29-JUN-1999; 99US-0140991P.
PR	30-JUN-1999; 99US-0141287P.

PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145911P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147030P.
PR	04-AUG-1999;	99US-0147204P.
PR	05-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	23-SEP-1999;	99US-0155139P.
PR	24-SEP-1999;	99US-0155486P.
PR	28-SEP-1999;	99US-0155659P.
PR	29-SEP-1999;	99US-0156458P.
PR	04-OCT-1999;	99US-0156596P.
PR	05-OCT-1999;	99US-0157117P.
PR	06-OCT-1999;	99US-0157753P.
PR	07-OCT-1999;	99US-0157865P.
PR	08-OCT-1999;	99US-0158029P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	22-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Alignment Scores:		
Pred. No.:	1.89e-79	816
Score:	768.00	149
Percent Similarity:	71.04%	Conservative: 35
Best Local Similarity:	57.53%	Mismatches: 73
Query Match:	56.14%	Indels: 2
DB:	3	Gaps: 2
US-10-780-002-2 (1-260) x AAC42344 (1-816)		
Qy	2	LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db	37	AAGAGGAACAACATTTGTACTAGTACATGTTGCGTCCATGGCGGTGGTGGTAC 96
Qy	22	LysLeuLysProLeuLeuGluAlaAalGlyHisLysValThrAlaLeuAspLeuAla 41
Db	97	AAGTTTAAGCGCTGTAGAGCGGTGGGCCCGCGTAACCTGCTGTGGACTTAGCTGCC 156
Qy	42	SerGlyThrAsp---LeuAtqLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db	157	TCCGGATATAGACAAACAGAGGTGCGATCCTGACATCCCCCATCGGAACATACTCGGAG 216
Qy	61	ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db	217	CCATTGACGAAGCTCCTGACCTCATTGCCAAATGATGAAAGGTTGTGCTCGTTGGTCAC 276
Qy	81	SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db	277	AGCTTTGGTGGCTTGAACTTAGCCATGCCATGAAAAGTTTCCCGAAAAAATCTCTGTC 336
Qy	101	AlaValPheLeuAlaAlaPheMetProAspSerValHisAanSerPheValLeuGlu 120

Db 337 GCTGTATCTTCTGCTTCTCATCGCGACACCGAACACTCCACCATCTCTGCTCTGGAC 396  
 Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 Db 397 AGTTTGGAGAACACATGCTTCAAGAAAGCATGGATGGCACCAGAAATCGAACCTTATGGT 456  
 Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
 Db 457 TCAGACAATCCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAGTTGGGTCTC 513  
 Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
 Db 514 TACCAGCTTCTCCAGTTCAGGACTTGAACCTGGGATTACTTTTAATGAGGCGCAGATCG 573  
 Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 Db 574 TTATTTATTAAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGT 633  
 Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 Db 634 CCTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATCCAGAGAACGCCAGAGATGG 693  
 Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
 Db 694 ATGATTGATAAATTTCCGGTGAATTTAGTGATGGAGATGGAGACAGATCATATGCCA 753  
 Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 Db 754 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTCGAATAATCCGCGACAATTC 810

## RESULT 4

ID ADA26530  
 XX ADA26530 standard; DNA; 918 BP.  
 AC ADA26530;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana AtSB2L1 gene homologous to tobacco SABP2 gene.  
 KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
 KW modulator; tobacco; ds; gene.  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2003016551-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 16-AUG-2002; 2002WO-US026312.  
 XX  
 PR 16-AUG-2001; 2001US-0312863P.  
 XX  
 PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 XX  
 PI Klessig DF, Kumar D;  
 XX  
 XX WPI; 2003-278578/27.

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.

PS Claim 65; Page: 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding  
 CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L1 which  
 CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
 CC (Note: this sequence is not given in the specification but is derived  
 CC from the Genbank accession number given by the inventors).

XX  
 SQ Sequence 918 BP; 266 A; 175 C; 221 G; 256 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	2,23e-79	Length:	918
Score:	768.00	Matches:	149
Percent Similarity:	71.04%	Conservative:	35
Best Local Similarity:	57.53%	Mismatches:	73
Query Match:	56.14%	Indels:	2
DB:	8	Gaps:	2

US-10-780-002-2 (1-260) x ADA26530 (1-918)

Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyr 21  
 Db 56 AAGAGGAACAACATTTTGTACTAGTACATGTTCTGTCATGCGCGCTGGTGTGGTAC 115  
 Qy 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
 Db 116 AAGTTAAGCGCTGCTAGAGCGGTGGCGCCACCGCTAACTGCTGTGACTAGTGC 175  
 Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
 Db 176 TCCGGATAGACACACGAGGTGCGATCTGACATCCCCACATGCGAACATATCTCGGAG 235  
 Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
 Db 236 CCATTGACGAAGCTCCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTCGTTGGTCA 295  
 Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 Db 296 AGCTTTGGTGGCTTGAACCTTAGCCATGCCATGAAAGTTTCCCGAAAAAATCTCTGTC 355  
 Qy 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
 Db 356 GCTGTATTCTTGACTGCTTTCATGCGGACACCCGACACTCACCATCTCTGCTTGGAC 415  
 Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 Db 416 AAGTTTGGAGCAACATGCTTCAAGAGCATGGATGGCCACCGAATTCGAACCTTATGGT 475  
 Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
 Db 476 TCAGACAATCCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAGTTGGGTCTC 532  
 Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
 Db 533 TACCAGCTTCTCCAGTTGAGATCTTGAACCTGGGATTACTTTTAATGAGGCCAGGATCG 592  
 Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 Db 593 TTATTTATTAAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGT 652  
 Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 Db 653 CCTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATTCAGAGAAACGCCAGAGATGG 712  
 Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
 Db 713 ATGATTGATAAATTTCCGGTGAATTTAGTGATGGAGATGGAGGACAGATCATATGCCA 772  
 Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259

Db	773	ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAAATCGCGGACAAATTC	829
RESULT 5			
AAC34411			
ID	AAC34411	standard; DNA; 870 BP.	
XX			
AC	AAC34411;		
XX			
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 6557.	
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999;	99US-0121825P.	99US-0139457P.
PR	05-MAR-1999;	99US-0123180P.	99US-0139458P.
PR	09-MAR-1999;	99US-0123548P.	99US-0139459P.
PR	23-MAR-1999;	99US-0125788P.	99US-0139460P.
PR	25-MAR-1999;	99US-0126264P.	99US-0139461P.
PR	29-MAR-1999;	99US-0126785P.	99US-0139462P.
PR	01-APR-1999;	99US-0127462P.	99US-0139463P.
PR	06-APR-1999;	99US-0128234P.	99US-0139750P.
PR	08-APR-1999;	99US-0128714P.	99US-0139763P.
PR	16-APR-1999;	99US-0129845P.	99US-0139817P.
PR	19-APR-1999;	99US-0130077P.	99US-0139899P.
PR	21-APR-1999;	99US-0130449P.	99US-0140354P.
PR	23-APR-1999;	99US-0130510P.	99US-0140695P.
PR	28-APR-1999;	99US-0130891P.	99US-0140823P.
PR	28-APR-1999;	99US-0131449P.	99US-0140991P.
PR	30-APR-1999;	99US-0132048P.	99US-0141287P.
PR	30-APR-1999;	99US-0132407P.	99US-0141842P.
PR	04-MAY-1999;	99US-0132484P.	99US-0142154P.
PR	06-MAY-1999;	99US-0132485P.	99US-0142055P.
PR	06-MAY-1999;	99US-0132486P.	99US-0142390P.
PR	06-MAY-1999;	99US-0132487P.	99US-0142803P.
PR	07-MAY-1999;	99US-0132863P.	99US-0142920P.
PR	11-MAY-1999;	99US-0134256P.	99US-0142977P.
PR	14-MAY-1999;	99US-0134218P.	99US-0143542P.
PR	14-MAY-1999;	99US-0134219P.	99US-0143624P.
PR	14-MAY-1999;	99US-0134221P.	99US-0144005P.
PR	18-MAY-1999;	99US-0134370P.	99US-0144085P.
PR	18-MAY-1999;	99US-0134768P.	99US-0144325P.
PR	19-MAY-1999;	99US-0134941P.	99US-0144331P.
PR	20-MAY-1999;	99US-0135124P.	99US-0144332P.
PR	21-MAY-1999;	99US-0135353P.	99US-0144333P.
PR	21-MAY-1999;	99US-0135629P.	99US-0144334P.
PR	25-MAY-1999;	99US-0136021P.	99US-0144335P.
PR	27-MAY-1999;	99US-0136392P.	99US-0144352P.
PR	28-MAY-1999;	99US-0136782P.	99US-0144632P.
PR	01-JUN-1999;	99US-0137222P.	99US-0144884P.
PR	03-JUN-1999;	99US-0137528P.	99US-0144814P.
PR	04-JUN-1999;	99US-0137502P.	99US-0145086P.
PR	07-JUN-1999;	99US-0137724P.	99US-0145088P.
PR	08-JUN-1999;	99US-0138094P.	99US-0145085P.
PR	10-JUN-1999;	99US-0138540P.	99US-0145087P.
PR	10-JUN-1999;	99US-0138847P.	99US-0145089P.
PR	14-JUN-1999;	99US-0139119P.	99US-0145192P.
PR	16-JUN-1999;	99US-0139452P.	99US-0145145P.
PR	16-JUN-1999;	99US-0139453P.	99US-0145218P.
PR	17-JUN-1999;	99US-0139454P.	99US-0145224P.
PR	17-JUN-1999;	99US-0139455P.	99US-0145267P.
PR	18-JUN-1999;	99US-0139456P.	99US-0145913P.
PR	18-JUN-1999;	99US-0139457P.	99US-0145918P.
PR	18-JUN-1999;	99US-0139458P.	99US-0145919P.
PR	18-JUN-1999;	99US-0139459P.	99US-0145951P.
PR	18-JUN-1999;	99US-0146386P.	99US-0145951P.
PR	18-JUN-1999;	99US-0146388P.	99US-0146389P.
PR	18-JUN-1999;	99US-0147038P.	99US-0147038P.
PR	18-JUN-1999;	99US-0147039P.	99US-0147039P.
PR	18-JUN-1999;	99US-0147040P.	99US-0147040P.
PR	18-JUN-1999;	99US-0147041P.	99US-0147041P.
PR	18-JUN-1999;	99US-0147042P.	99US-0147042P.
PR	18-JUN-1999;	99US	

PR	20-AUG-1999;	99US-0149722P.	QY	22	LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla	41
PR	20-AUG-1999;	99US-0149723P.	Db			
PR	20-AUG-1999;	99US-0149929P.		93	AAGGTTAAGCCTTCTTCGAGGCTTTGGCCATCGTGTAACCGCTTAGACCTAGCTGCT	152
PR	23-AUG-1999;	99US-0149902P.				
PR	23-AUG-1999;	99US-0149930P.	QY	42	SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu	60
PR	25-AUG-1999;	99US-0150566P.	Db			
PR	26-AUG-1999;	99US-0150884P.		153	TCCGGTATAGACACCAACCGGTCATCACTGACATTTCTACATGTGAACAATATTCTGAG	212
PR	27-AUG-1999;	99US-0151065P.	QY	61	ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis	80
PR	27-AUG-1999;	99US-0151080P.	Db			
PR	30-AUG-1999;	99US-0151303P.		213	CCATTGATGCAGCTAATGACTTTCATTCGCGAATGATGAGAGGTGTACTCGTTGGTCA	272
PR	31-AUG-1999;	99US-0151438P.	QY	81	SerLeuGlyGlyMetAsnLeuAlaMetGluLysTyrProGlnLysIleTyrAla	100
PR	01-SEP-1999;	99US-0151930P.	Db			
PR	07-SEP-1999;	99US-0152363P.		273	AGCTTTGGAGGTTTGAGTTTAGCCTTAGCCATGGATAAGTTTCCCGATAAATCTCTGC	332
PR	10-SEP-1999;	99US-0153070P.	QY	101	AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu	120
PR	13-SEP-1999;	99US-0153758P.	Db			
PR	15-SEP-1999;	99US-0154018P.	QY	121	GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly	140
PR	16-SEP-1999;	99US-0154039P.	Db			
PR	20-SEP-1999;	99US-0154779P.		393	AGTTTGAAGCAGCATGACACAGAGATGGATGGGCTCTGAGCTCGACACATATGCT	452
PR	22-SEP-1999;	99US-0155139P.	QY	141	SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu	160
PR	23-SEP-1999;	99US-0155486P.	Db			
PR	24-SEP-1999;	99US-0155659P.		453	TCAGATAATTCGGGCTTG---TCTGTGTTCTTCAGCACCAGCTTCATGAAGCACCGTCTC	509
PR	28-SEP-1999;	99US-0156458P.	QY	161	TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer	180
PR	29-SEP-1999;	99US-0156596P.	Db			
PR	04-OCT-1999;	99US-0157117P.		510	TACCAACTTCTCCTGTGGAGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTAGTTCA	569
PR	05-OCT-1999;	99US-0157753P.	QY	181	LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal	200
PR	06-OCT-1999;	99US-0157865P.	Db			
PR	07-OCT-1999;	99US-0158029P.		570	TGTTTATTAATGAATATCGAAGATGGAGAACTTTCTGAAAGGGTATGGATCTGTT	629
PR	08-OCT-1999;	99US-0158232P.	QY	201	LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp	220
PR	12-OCT-1999;	99US-0158369P.	Db			
PR	13-OCT-1999;	99US-0159293P.		630	CCTCGAGCTTACATTGTGTGCAAGAGAGCAACATTATCTCGGAAGACCATCAACGATGG	689
PR	13-OCT-1999;	99US-0159294P.	QY	221	GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla	240
PR	14-OCT-1999;	99US-0159329P.	Db			
PR	14-OCT-1999;	99US-0159331P.		690	ATGATCCATAATTATCCGGCGAATTTAGTGATTGAGATGGAAGAGACGATCATATGCCA	749
PR	14-OCT-1999;	99US-0159637P.	QY	241	MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn	260
PR	18-OCT-1999;	99US-0159638P.	Db			
PR	21-OCT-1999;	99US-0160741P.		750	ATGTTTGCACAACTCAAGTACTAAGTACCATCTATTGGCAATCGCTGACAAATTTCTCT	809
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161933P.				
PR	29-OCT-1999;	99US-0162142P.				

## Alignment Scores:

Pred. No.: 5,08e-71 Length: 870  
 Score: 696.00 Matches: 138  
 Percent Similarity: 67.31% Conservative: 37  
 Best Local Similarity: 53.08% Mismatches: 83  
 Query Match: 50.88% Indels: 2  
 DB: 3 Gaps: 2

US-10-780-002-2 (1-260) x AAC34411 (1-870)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyr 21  
 ||| :|||  
 Db 33 AAGAGGAAGCAACACTTCGTGTAGTACATGGTGGCGGCATGGTGGTAC 92

```
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1551; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid with an
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 792 BP; 210 A; 169 C; 187 G; 226 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5,81e-71 Length: 792
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 6 Gaps: 2
US-10-780-002-2 (1-260) x AB213746 (1-792)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
DB 13 AAGAGGAAGCAACACTTCGTGTAGTACATGCTGGTCCGCGCATGGTGGTGGTAC 72
QY 22 LysLeuLysProLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
DB 73 AAGTTAAGCCTCTTCGAGGCTTGGCCATCGTGTAACCGCTTAGACCTAGCTGCT 132
QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
DB 133 TCCGGTATAGACACACACAGGTCATCTACATGACATTTCTACATGGAACAATATCTGAG 192
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 193 CCATTGATGCGCTAATGACTTCATTGCCGAATGATGAGAGGTTGTACTCGTGGTCAAT 252
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 253 AGCTTTGGAGGTTTGAGTTAGCTTAGCCATGATGATAGTTCCCGATAAATCTCTGTC 312
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
DB 313 TCTGTCTTCGTGACTGCAATTCATGCCGACACCAACACTCACCATCGTTCGTCGAGGAA 372
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
DB 373 AAGTTTGCAGGAGCATGACACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 432
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
DB 433 TCAGATAAATCCGGCTTG---TCTGTGTCTTCAGCACCAGCTTCATGAAGCACCGCTCTC 489
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
DB 490 TACCAACTTTCCTCTGTGGAGGATCTTGAGCTTGGATTGCTTAAAGAGGCGCTAGTTCA 549
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 550 TTGTTTATTAATGAATTATCGAATCGAGAACCTTTTCTGAGAAAGGATGATCTGTT 609
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 610 CCTCGAGCTTACATTGTGTGCAAAAGAGGACAACTTATCTCGAAGACCATCAACGATGG 669
QY 221 GlnIleAspAnlleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 670 ATGATCCATAATTATCCGGCGAATTTAGTATTGAGATGGAAGAGACTGATCATATGCCA 729
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuGluLeuAlaHisLysTyr 259
Db 730 ATGTTTGCAAAACCTCAACTACTAGTGACCATCTATTGGCAATCGCTGACAAATTTC 786
RESULT 7
AAC47519
ID AAC47519 standard; DNA; 864 BP.
XX
AC AAC47519;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 54118.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
```



PR 03-JUN-1999;	99US-01337528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 05-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 20-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		

Alignment Scores:

```
Pred. No.: 6,58e-71 Length: 864
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 3 Gaps: 2

US-10-780-002-2 (1-260) x AAC47519 (1-864)

QY 2 LysGluGlyLysPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
DB 31 AAGAGGAGCAACACTTCGTGTAGTACATGCTGCGCCACGGCGCATGGTGGTAC 90

QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
DB 91 AAGGTTAAGCCTCTTCTCAGGCTTTGGGCCATCGTGAACCGCTTAGACCTAGCTGCT 150

QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyAspThrLeu 60
DB 151 TCCGGTATAGACCAACACGAGTCAATCTGACATCTTCTACATGTGAACAATATTCTGAG 210

QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 211 CCATTGATGACGCTAATGACTTCATTCGCGAATGATGAGAGGTGTACTCGTGTGTCAT 270

QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 271 AGCTTTGGAGGTTTGAGTTTAGCTTAGCCATGATGAAGTTTCCCGATAAAATCTCTGTC 330

QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGlu 120
DB 331 TCTGTCTTCGTGACTGCAATTCGCGCACCAACACTCACCATCGTTCGTCGAGGAA 390

QY 121 GlyTrpAsnGluArgThrProAlaGluAenTrpLeuAspThrGlnPheLeuProTyrGly 140
DB 391 AGTTTGCAGCAGCATGACACAGAGATGGGCTCTGAGCTCGAGACATATGTT 450

QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
DB 451 TCAGATAAATTCGGGCTTG---TCTGTGTCTTTCAGCACCAGCACTTCATGAAGCACCCTCTC 507

QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
DB 508 TACCACCTTCTCCTGTGAGGATCTTGAGCTTGATGCTTCTTAAGAGGCTAGTTCA 567

QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
DB 568 TTGTTTATTAATGAATATTCGAGATGAGAGACTTTTCTGAGAAAGGGTATGATCTGTT 627

QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
DB 628 CCTCGAGCTTACATTGTGTGCAAGAGGACAACTTATCTCGAAGACCATCAACGATGG 687

QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
DB 688 ATGATCCATAATTTATCCGCGCAATTTAGTGTGATGTGAGATGGAAGAGACTCATATGCCA 747

QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 748 ATGTTTTCGAACCTCACTACTACTAGTGACCATCTATTGGCAATCGCTGACAAATTTC 804

RESULT 8
ADA26531
ID ADA26531 standard; DNA; 950 BP.
XX
AC ADA26531;
XX
XX 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana AtSB2L2 gene homologous to tobacco SABP2 gene.
XX
KW salicylic acid-binding protein; SABP2; disease resistance; plant;
modulator; tobacco; db; gene.
XX
```

```
XX Arabidopsis thaliana.
OS
XX WO2003016551-A2.
PN
XX 27-FEB-2003.
PD
XX
XX 16-AUG-2002; 2002WO-US026312.
PF
XX 16-AUG-2001; 2001US-0312863P.
PR
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
PA
XX Klessig DF, Kumar D;
PI
XX WPI; 2003-278578/27.
DR
XX
XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for
PT producing transgenic plants with increased resistance to disease, or for
PT screening SABP2 modulators that confer enhanced resistance of plants to
PT disease.
XX
XX Claim 65; Page; 98pp; English.
XX
XX The invention relates to a novel isolated 1079 bp salicylic acid-binding
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid
CC protein, sequences that specifically hybridize with the nucleic acid, the
CC complement of the nucleic acid or a natural allelic variant of the
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for
CC conferring increased resistance to disease in plants, or for screening
CC modulators of SABP2, which confer increased or enhanced resistance of
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also
CC useful for producing transgenic plants with increased resistance to
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful
CC for studying the molecular mechanisms responsible for the modulation of
CC salicylic acid-mediated disease resistance in plants. This sequence
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L2 which
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.
CC (Note: this sequence is not given in the specification but is derived
CC from the Genbank accession number given by the inventors).
XX
SQ Sequence 950 BP; 272 A; 185 C; 218 G; 275 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,54e-71 Length: 950
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x ADA26531 (1-950)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
DB 45 AAGAGGAGCAACACTTCGTGTAGTACATGCTGCGCCACGGCGCATGGTGGTAC 104

QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
DB 105 AAGGTTAAGCCTCTTCTCAGGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCT 164

QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyAspThrLeu 60
DB 165 TCCGGTATAGACCAACACGAGTCAATCTGACATCTTCTACATGTGAACAATATTCTGAG 224

QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 225 CCATTGATGACGCTAATGACTTCATTGCGCAATGATGAGAGGTGTACTCGTGTGTCAT 284

QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 285 AGCTTTGGAGGTTTGAGTTTAGCTTAGCCATGATGAAGTTTCCCGATAAAATCTCTGTC 344
```

```
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
    ::::::::::: |||||::: ||::: ||::: |||||::: |||
Db 345 TCTGCTCTGAGCTGACATTATGCGCGACCAACACTCACCATCGTTGTCGAGGAA 404
    ::::::::::: |||||::: ||::: ||::: |||||::: |||
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
    ::::::::::: |||||::: ||::: ||::: |||||::: |||
Db 405 AAGTTTTCGAAGCAGCATGACACAGAAAGGATGGCTCTGAGCTCGAGACATATGTT 464
    ::::::::::: |||||::: |||||::: |||||::: |||||::: |||
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
    ||| ||| |||::: |||||::: |||||::: |||||::: |||||::: |||
Db 465 TCAGATAATTCGGCTTG---TCTGTGTTCTTACGACCGACTTCAAGACACCGCTCTC 521
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
Db 522 TACCAACTTCTCCTGTGGAGGATCTTGACITTTGGATTGCTTAAAGAGGCTAGTTCA 581
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
Db 582 TTGTTTATTAAATGAATATCGAAGATCGAAGAACTTTTCTGAGAAAGGATATGATCTGTT 641
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTTrp 220
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
Db 642 CCTCGAGCTTACATTGTGTGCAAGAGGAAACAATTATCTCGAAGACCATCAACATGG 701
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
Db 702 ATGATCCATAATTCGCGGGAATTTAGTGATTGAGATGGAAGAGACTCATATGCCA 761
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||
Db 762 ATGTTTTGCAAACTCAACTACTAAGTGACCATCTATTGGCAATCGCTGACAATTC 818
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||

RESULT 9
AAC43133
ID AAC43133 standard; DNA; 792 BP.
AC AAC43133;
XX
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38151.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 99US-0145087P.
PR 99US-0145087P.
```



QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluAlaHisLysTyr 259  
 DB 739 AAGCCTCAACAACTATGTCTCTTCTTGGAGATTGCAATATAT 786

## RESULT 10

ID AB212412  
 AC AB212412 standard; DNA; 792 BP.

XX AB212412;

XX 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 217.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264547P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and

PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 217; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office

XX Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,22e-70 Length: 792  
 Score: 690.00 Matches: 134  
 Percent Similarity: 69.92% Conservative: 45  
 Best Local Similarity: 52.34% Mismatches: 75  
 Query Match: 50.44% Indels: 2  
 DB: 6 Gaps: 2

US-10-780-002-2 (1-260) x AB212412 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrLysLeuLys 24

DB 22 CACACGTCGTTCTAGTATAGTGTGCTTCCATGGCGCTGTGCTGACAGGTAA 81

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44

DB 82 CCGCAGCTCGAGGCTTCTGGCCACCGCGTAACCGCGGTAGATCTAGCTGCTCCGATATA 141

QY 45 AspLeu---ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63  
 DB 142 GACATGACCAGGTCAATCACAGATATATCCACATCGGAACATATCTCAGAGCCATTGATG 201

QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83  
 DB 202 CAGCTAATGACCTCACTACCGATGATGAGAAGGTTGTCTTGTGGTCATAGCTTAGGA 261

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103  
 DB 262 GGTITGAGTTAGCTATGCGCATGATGTTCCGACCAAAATCTCTGTTTCTGTCTTT 321

QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123  
 DB 322 GTGACTGCTATGATGCCAGACACCAACACCTCCTCGTATGGATAAGCTAAGA 381

QY 124 GluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGlySerProGlu 143  
 DB 382 AAAGAAACTTTCACGAGAGGAATGGTTAGACACCGTGTTCGAGC---GAGAAACCTGAT 438

QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163  
 DB 439 TTTCTCTAGCGAGTTTGTGATTTTTCGACCAAGATTTCATGGCCCAAGAACTTGTATCAGTTG 498

QY 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183  
 DB 499 TCTCAGTCCAGATCTTGAATTCGGAATAATGTTGGTGGGCAACCCATTGATTAG 558

QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203  
 DB 559 AAAGATATGGCAGAGAGAAGAGCTTCAGTGAGGAAGGATACCGATCCGTTACACGTATA 618

QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrpGlnIleAsp 223  
 DB 619 TTTATTGTATGCGGAAGGATCTTGTGTCCACCCGAAGATTACCGCATCGATGATCAGC 678

QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 DB 679 AACTTTCCCAAAAGAGTAATGAGATCAAGACGCGCATCATATGCCAATGTTCTCC 738

QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluAlaHisLysTyr 259

DB 739 AAGCCTCAACAACTATGTCTCTTCTTGGAGATTGCAATATAT 786

## RESULT 11

ID ADN74620 standard; cDNA; 792 BP.

XX AC ADN74620;

XX DT 15-JUL-2004 (first entry)

XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2515.

XX gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

XX growth regulator; animal feed product; thale cress;

XX cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX

DR WPI; 2004-348466/32.  
DR P-PSDB; ADN74621.  
XX  
PT Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprises modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.  
XX  
PS Claim 1; SEQ ID NO 2515; 134pp; English.  
XX  
CC This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC BzFa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreplication, biochemistry, signal  
CC transduction, storage lipid mobilization and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is thale cress CDNA  
CC repressed 1.3 fold or more in plants overexpressing the BzFa/Dpa  
CC transcription factor, given in an exemplification of the invention.  
XX  
SQ Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,22e-70 Length: 792  
Score: 650.00 Matches: 134  
Percent Similarity: 69.92% Conservative: 45  
Best Local Similarity: 52.34% Mismatches: 75  
Query Match: 50.44% Indels: 2  
DB: 12 Gaps: 2

US-10-780-002-2 (1-260) x ADN74620 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyrIleuLys 24  
DB 22 CAACACGCTGCTAGTACATGGTCTGCCATCCCGCTGCTGGTGGTACAGGTAA 81  
QY 25 ProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44  
DB 82 CGCAGCTCGAGGCTTCTGGCCACCGCTAACCGCGTAGATCTAGTGCCTCCGGTATA 141  
QY 45 AspLeu---ArgGlyIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63  
DB 142 GACATGACCGAGTCAATCACAGATATATCCATCGCAACAATCTACAGAGCAATTGATG 201  
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83  
DB 202 CAGCTAATGACCTCACTACCATGATGAGAAAGGTGTGCTTGTGGTCATAGCTTAGGA 261  
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103  
DB 262 GGTGTGAGTTAGCTATGCGCCATGGATATGTTTCCGACCAAAATCTCTGTTCTGCTTT 321  
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123  
DB 322 GTGACTGCTATGATGCCACACCAACACTCACCATCTCTGTATGGATAAGCTAAGA 381  
QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143  
DB 382 AAAGAAACTTCACGAGGAATGGTTAGACACCGTGTTCACGAGC---GAGAAACCTGAT 438  
QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163  
DB 439 TTTCCTAGCGAGTTTGGATTTTGGACCAGAAATTCATGCCCAAGAACTTGTATCAGTTG 498

QY 164 CysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183  
DB 499 TCTCAGTCCAGATCTTGAATTCGCGAAAATGTTGGTGGAGGCAAAACCCATTGATTAA 558  
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValIleArgVal 203  
DB 559 AAAGATATGGCAGAGAGAAAGAGCTTCAGTGAGGAGGATACGGATCCGTTACAGGTATA 618  
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223  
DB 619 TTTATTGATTCGGAAGAGATCTTGTGTCCACCGAAGATTACCAGCGATCGATGATCAGC 678  
QY 224 AsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAlaMetLeuCys 243  
DB 679 AACTTTCCTCCCAAGAGTAAATGAGATCAAAAGCCGAGATCATATGCCAATGTTCTCC 738  
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
DB 739 AAGCCTCAACAATATGTCTTCTCTTGAGATTGCAAAATAAATAT 786  
RESULT 12  
ADA26532  
ID ADA26532 standard; DNA; 1152 BP.  
XX  
AC ADA26532;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana AtSB2L3 gene homologous to tobacco SABP2 gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.  
XX  
OS Arabidopsis thaliana.  
XX  
FN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
PI Klessig DF, Kumar D;  
XX  
DR WPI; 2003-278578/27.  
XX  
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.  
XX  
PS Claim 65; Page; 98pp; English.  
XX  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding  
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L3 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).

```
XX SQ Sequence 1152 BP; 372 A; 214 C; 253 G; 313 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.8e-70 Length: 1152
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservativity: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x ADA26532 (1-1152)
QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLeuLeuLys 24
DB 151 CAACAGCTGCTTCTAGTACATGGTGTTCCTGCGCCCTGGTGTGGTACAGGTTAG 210
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44
DB 211 CCGCAGCTCGAGGCTTCTGGCCACCGCTAACCGCGTAGATCTAGCTCGCTCCGGTATA 270
QY 45 AspLeu---ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
DB 271 GACATGACGAGGTCATACAGATATATCCACATGCGAACAATACTCAGAGCCATTGATG 330
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
DB 331 CAGCTAATGACCTCATCACCAGATGATGAGAGGTGTGCTGTGTGTCATAGCTTAGGA 390
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
DB 391 GGTTCAGTTAGTATGCTGCGATGATATGTTCCGACCAAAATCTCTGTTCTGCTTT 450
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
DB 451 GTGACTGCTATGATGCCAGACACCAACACCTCACCATCTCTGATGGGATAAGCTAAGA 510
QY 124 GluArgThrProAlaGlnAsnTyrLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
DB 511 AAAGAACTTCAGGAGAGAAATGTTAGACCGTGTTTACAGC---GAGAAACCTGAT 567
QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
DB 568 TTTCTCTAGCGAGTTTGGATTTTGGACCAAGAATTCATGCCCAAGAACTTGTATCAGTTG 627
QY 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
DB 628 TCTCCAGTCCAGATCTTGAATTTGGCGAAAATGTTGGTGAGGGCAAAACCATTTAAG 687
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
DB 688 AAAGATATGCGAGAGAGAGAGCTTCAGTGAGGAGGATACGGATCCGTTACAGTATA 747
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyrGlnIleAsp 223
DB 748 TTTATTGTATGCGAAGGATCTGTGTCAACCGAAGATTACCAGCATCGATGATCAGC 807
QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
DB 808 AACTTTTCCCCAAAAGAAAGTATGGAGATCAAAGACGCGAGATCATATGCGCAATGTCTCC 867
QY 244 GluProGlnLysLysCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 868 AAGCTTCAAACTATGTGCTCTTCTTTGGAGATTGCAATATAATAT 915

RESULT 13
ID ADA26533
XX ADA26533 standard; DNA; 1109 BP.
XX AC ADA26533;
XX DT 20-NOV-2003 (first entry)
XX
```

DE Arabidopsis thaliana AtSB2L4 gene homologous to tobacco SABP2 gene.

XX salicylic acid-binding protein; SABP2; disease resistance; plant;

KW modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

OS WO2003016551-A2.

PN 27-FEB-2003.

PD 16-AUG-2002; 2002WO-US026312.

PF 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PA Klessig DF, Kumar D;

PI WPI; 2003-278578/27.

DR New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for

PT producing transgenic plants with increased resistance to disease, or for

PT screening SABP2 modulators that confer enhanced resistance of plants to

PT disease.

XX Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid

CC protein, sequences that specifically hybridize with the nucleic acid, the

CC complement of the nucleic acid or a natural allelic variant of the

CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for

CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also

CC useful for producing transgenic plants with increased resistance to

CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful

CC for studying the molecular mechanisms responsible for the modulation of

CC salicylic acid-mediated disease resistance in plants. This sequence

CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L4 which

CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.

CC (Note: this sequence is not given in the specification but is derived

CC from the Genbank accession number given by the inventors).

XX SQ Sequence 1109 BP; 365 A; 193 C; 235 G; 316 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.61e-69 Length: 1109

Score: 680.50 Matches: 134

Percent Similarity: 67.58% Conservativity: 39

Best Local Similarity: 52.34% Mismatches: 82

Query Match: 49.74% Indels: 1

DB: 8 Gaps: 1

US-10-780-002-2 (1-260) x ADA26533 (1-1109)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLeuLys 24

DB 19 AAGCGGTTTGTCTCGTCCATGGGTGTGCCACGCGCTTGTGTGTGTACAAAGGTGAA 78

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44

DB 79 ACGCATCTGGAGGCTGTAGTCTGTGTGACCGCGGTGGATCTAGCTGCAATCCGGTATA 138

QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64

DB 139 AATATGACTAGATTGGAAGAGATTTCAGACTTTGAAGGATTACTGCAACACCTTTGCTTGA 198

QY 65 LeuMetGluSerLeu---SerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83

DB 199 TTACTGAACCTCACTTGGCTCGGATGACGATGAAGGTGATTCTTTGTCGCATAGTAGTGA 258

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103  
 DB 259 GGAATACCTGCTGCTCTGACATATTCCTAGTAAGATTGCTACTATTGTTTTC 318  
 QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123  
 DB 319 TTGACAGCTTTTATGCCGACACAGAACCTACCTGCTTATGTTTACCAAGCTAATC 378  
 QY 124 GluA-gThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143  
 DB 379 AGAAGCGTTTCCACAGAAGAGTGTGGACACCGTGTGGAAACGTATGGGAACATGAA 438  
 QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163  
 DB 439 TGTCTCTAGAGTTTGCTCTTTTGGACCAAGTTTCATGGCCCAAGAAATTTGTATCACTC 498  
 QY 164 CysSerProGluAspLeuAlaLaserSerLeuValArgProSerSerLeuPheMet 183  
 DB 499 TCTCCGTCACAGATCTTGAATTGGGAAAATGTTGGTGAGAGTAACCCCATATTACA 558  
 QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203  
 DB 559 AATAATCTGGCAGGACAAAGAGCTTTAGTGAGGAAGGTACCGTTACCGTTACAGCTATA 618  
 QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223  
 DB 619 TATATTGTATGTGGAGAGACATGCGGTACCCGAGGATTACCACTGGTGGATGATCAAG 678  
 QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 DB 679 AACTTCCGCCAAAGAGTATGAGATCAATGTGCAGATCATATGGCAATGTTCTCC 738  
 QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 DB 739 AAGCCTCACAACTATGTCTCTCTCGTGGAGATTGCGATGTAATAT 786  
 RESULT 14  
 ADA26535  
 ID ADA26535 standard; DNA; 1046 BP.  
 XX  
 AC ADA26535;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana AtSB2L6 gene homologous to tobacco SABP2 gene.  
 XX  
 KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
 KW modulator; tobacco; ds; gene.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W02003016551-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 16-AUG-2002; 2002WO-US026312.  
 XX  
 PR 16-AUG-2001; 2001US-0312863P.  
 XX  
 PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 XX  
 PI Klessig DF, Kumar D;  
 XX  
 DR WPI; 2003-278578/27.  
 XX  
 PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.  
 XX  
 PS Claim 65; Page; 98pp; English.  
 XX  
 CC The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening  
 CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L6 which  
 CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
 CC (Note: this sequence is not given in the specification but is derived  
 CC from the Genbank accession number given by the inventors).  
 XX

SQ Sequence 1046 BP; 330 A; 195 C; 248 G; 273 T; 0 U; 0 Other;

Alignment Scores: 2.37e-67 Length: 1046  
 Pred. No.: 665.50 Matches: 128  
 Score: 67.57% Conservative: 47  
 Percent Similarity: 49.42% Mismatches: 83  
 Best Local Similarity: 48.65% Indels: 1  
 Query Match: 8 Gaps: 1  
 DB:

US-10-780-002-2 (1-260) x ADA26535 (1-1046)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyr 21  
 DB 31 AAGAACCAAGACGGTTTGTCTCATCCATGGGTGGCCACGGGGCATGACTTGGGAC 90  
 QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
 DB 91 AAGGTGAAAACGACGCTGGAGTTGCAGTTCACATGTGTGACGCGAGTGTCTTGTGCA 150  
 QY 42 SerGlyThrAspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
 DB 151 TCAGTATAACATGACCCAAAGTGGAGAGATTTCAGACTCTGAACGATTACTGCAACCA 210  
 QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGlu---LysValIleLeuValGlyHis 80  
 DB 211 TTGCTTGAGTTCTGAGCTCGCTTGGCTCAGATGACGGTAAGTGTGTTGTTGCTCAT 270  
 QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 DB 271 AGCATGGGAGGAATATCCGCTGCACTTCTGCTGACAGCTTCGCTTGAAGATTGCCGCT 330  
 QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
 DB 331 ATTGTCTTTTGGACAGCTTTCATGCCGACACAATAAACCCACCTGCTTATGTTTACGAA 390  
 QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 DB 391 AAGCTGCTCAGACGATTCACAGAGGAATGTTGGACACACCGCTGTGTGAACATACGG 450  
 QY 141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160  
 DB 451 AAACCTGATTTTCTCTACAGTATATCTTTTGGGACCAAAAGTTTATGCCCAAGAAATG 510  
 QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLaserSerLeuValArgProSerSer 180  
 DB 511 TATCAAAACTCTCCAGTCTTCAAGATCTTGAAGTGGTGAAGACATTAGTAGGGGAAACCCG 570  
 QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 DB 571 TTAGTTACAAACAATCTGGCAGGACAAAGAGCTTTAGTAGGAAGGCTACGATCCGCTT 630  
 QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 DB 631 ACACGTATATATATTGTATGACAGAGGATCTTGTGGAAAGTCGAAGATTACACGGTTGG 690  
 QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240



Db	691	ATGATTAGCAACTTTCACCAAGAAAGTAATGGAGATCAAATGTGCAGATCATATGCCA	750	PR	16-JUN-1999;	99US-0139453P.
		:::		PR	17-JUN-1999;	99US-0139492P.
		:::		PR	18-JUN-1999;	99US-0139454P.
Qy	241	MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr	259	PR	18-JUN-1999;	99US-0139455P.
		:::		PR	18-JUN-1999;	99US-0139456P.
Db	751	ATGTTCTCCAAGCCTCAAGAAGTTTGTGCTCTTCTCGAGATTGCAATAAATAT	807	PR	18-JUN-1999;	99US-0139457P.
		:::		PR	18-JUN-1999;	99US-0139458P.
		:::		PR	18-JUN-1999;	99US-0139459P.
		:::		PR	18-JUN-1999;	99US-0139460P.
		:::		PR	18-JUN-1999;	99US-0139461P.
		:::		PR	18-JUN-1999;	99US-0139462P.
		:::		PR	18-JUN-1999;	99US-0139463P.
		:::		PR	18-JUN-1999;	99US-0139750P.
		:::		PR	18-JUN-1999;	99US-0139763P.
		:::		PR	21-JUN-1999;	99US-0139817P.
		:::		PR	22-JUN-1999;	99US-0139899P.
		:::		PR	23-JUN-1999;	99US-0140353P.
		:::		PR	23-JUN-1999;	99US-0140354P.
		:::		PR	24-JUN-1999;	99US-0140695P.
		:::		PR	28-JUN-1999;	99US-0140823P.
		:::		PR	29-JUN-1999;	99US-0140991P.
		:::		PR	30-JUN-1999;	99US-0141287P.
		:::		PR	01-JUL-1999;	99US-0141842P.
		:::		PR	01-JUL-1999;	99US-0142154P.
		:::		PR	02-JUL-1999;	99US-0142055P.
		:::		PR	06-JUL-1999;	99US-0142390P.
		:::		PR	08-JUL-1999;	99US-0142803P.
		:::		PR	09-JUL-1999;	99US-0142920P.
		:::		PR	12-JUL-1999;	99US-0142977P.
		:::		PR	13-JUL-1999;	99US-0143542P.
		:::		PR	14-JUL-1999;	99US-0143624P.
		:::		PR	15-JUL-1999;	99US-0144005P.
		:::		PR	16-JUL-1999;	99US-0144085P.
		:::		PR	16-JUL-1999;	99US-0144086P.
		:::		PR	19-JUL-1999;	99US-0144325P.
		:::		PR	19-JUL-1999;	99US-0144331P.
		:::		PR	19-JUL-1999;	99US-0144332P.
		:::		PR	19-JUL-1999;	99US-0144333P.
		:::		PR	19-JUL-1999;	99US-0144334P.
		:::		PR	19-JUL-1999;	99US-0144335P.
		:::		PR	20-JUL-1999;	99US-0144352P.
		:::		PR	20-JUL-1999;	99US-0144632P.
		:::		PR	21-JUL-1999;	99US-0144884P.
		:::		PR	21-JUL-1999;	99US-0144814P.
		:::		PR	21-JUL-1999;	99US-0145086P.
		:::		PR	21-JUL-1999;	99US-0145088P.
		:::		PR	22-JUL-1999;	99US-0145085P.
		:::		PR	22-JUL-1999;	99US-0145087P.
		:::		PR	22-JUL-1999;	99US-0145089P.
		:::		PR	22-JUL-1999;	99US-0145192P.
		:::		PR	23-JUL-1999;	99US-0145145P.
		:::		PR	23-JUL-1999;	99US-0145218P.
		:::		PR	23-JUL-1999;	99US-0145224P.
		:::		PR	26-JUL-1999;	99US-0145276P.
		:::		PR	27-JUL-1999;	99US-0145913P.
		:::		PR	27-JUL-1999;	99US-0145918P.
		:::		PR	27-JUL-1999;	99US-0145919P.
		:::		PR	28-JUL-1999;	99US-0145951P.
		:::		PR	02-AUG-1999;	99US-0146386P.
		:::		PR	02-AUG-1999;	99US-0146388P.
		:::		PR	02-AUG-1999;	99US-0146389P.
		:::		PR	03-AUG-1999;	99US-0147038P.
		:::		PR	04-AUG-1999;	99US-0147204P.
		:::		PR	04-AUG-1999;	99US-0147302P.
		:::		PR	05-AUG-1999;	99US-0147192P.
		:::		PR	05-AUG-1999;	99US-0147260P.
		:::		PR	06-AUG-1999;	99US-0147303P.
		:::		PR	06-AUG-1999;	99US-0147416P.
		:::		PR	09-AUG-1999;	99US-0147493P.
		:::		PR	09-AUG-1999;	99US-0147935P.
		:::		PR	10-AUG-1999;	99US-0148171P.
		:::		PR	11-AUG-1999;	99US-0148319P.
		:::		PR	12-AUG-1999;	99US-0148341P.

RESULT 15

AAC49239  
ID AAC49239 standard; DNA; 984 BP.  
XX AAC49239;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60437.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EF1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000BP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.

```
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrTrpSerTyrTyrLysLeuLys 24
:::|||||
Db 22 CAACACGTGCTTAGTACATGGTGGCTTCATCGCCGCTGGTCTGGTCAAGGTTAAG 81
|||||
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
|||||
Db 82 CCGCAGCTCGAGGCTTCTGGCCACCGGTAAACCGCTAGATCTAGCTGCTCCGGTATA 141
|||||
QY 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
|||:::|
Db 142 GACATGACCGGTCAATCAGATATATCCATGCGAACAATCTCAGAGCCATTGATG 201
|||||
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
:::|||||
Db 202 CAGCTAATGACCTCACATGATGATGATGATGATGATGATGATGATGATGATGATG 261
|||||
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
|||||
Db 262 GGTGGAGTTAGCTATGGCCATGGATGATGATGATGATGATGATGATGATGATGATG 321
|||||
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheVal----- 118
:::|||||
Db 322 GTGACTGCTATGATGCCAGACACCAAACTCACCATCTTCGTATGGGATTAAGATGSCA 381
|||||
QY 118 ----- 118
Db 382 CTTGAATCATAACGCTAGGATACCAATTTTGAAGCTATGGCTTCTTATTCACACAT 441
|||||
QY 118 ----- 118
Db 442 AATGAGGGGTTTCATTTTCATGATTTTAGAACTACAAAAATTTCTATACGTGGAGTTGCT 501
|||||
QY 119 -----Leu 119
Db 502 CTTACTTACATAACTGTTTAAAAAATTTTACAAAAATGAAAAATAAACTACTTAATCTTATC 561
|||||
QY 120 GluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyr 139
:::|||||
Db 562 CAACAACCTAAGAAAGAAACTTCACGAGAGAAATGGTTAGACACCGTGTTCAGAGC--- 618
|||||
QY 140 GlySerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLys 159
|||||
Db 619 GAGAAACCTGATTTTCTAGCGAGTTTGGATTTTGGACCCAGATTCATGCGCAAGAAC 678
|||||
QY 160 LeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSer 179
|||||
Db 679 TTGTATCAGTTGTCTCCAGTCCAAAGATCTTGAATTGGCGAAAAATGTTGGTGGGCGCAAC 738
|||||
QY 180 SerLeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySer 199
|||||
Db 739 CCATTGATTAAGAAGATATGCGAGAGAGAAAGATCTTGTGTCCACCGAAGATTACCGATCC 798
|||||
QY 200 VallysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArg 219
|||||
Db 799 GTTACACGTATATTTATTTATGCGGAAGATCTTGTGTCCACCGAAGATTACCGCGCA 858
|||||
QY 220 TrpGlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMet 239
|||||
Db 859 TCGATGATCAGCAACTTTTCCCAAAAGAAAGTAAATGGAGATCAAAAGACGAGATCATATG 918
|||||
QY 240 AlaMetLeuCysGluProGlnLysCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
|||||
Db 919 CCAATGTTCTCCAAGCCTCAACAACTATGTGCTCTTCTTGGAGATTGCAATAAATAT 978
|||||
```

Search completed: September 12, 2005, 20:58:16

Job time : 491 secs

```
Alignment Scores:
Pred. No.: 2,78e-66
Score: 656.00
Length: 984
Percent Similarity: 56.25%
Best Local Similarity: 42.19%
Query Match: 47.95%
DB: 3
Matches: 135
Conservatives: 45
Mismatch: 74
Indels: 66
Gaps: 3
```

US-10-780-002-2 (1-260) x AAC49239 (1-984)

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 12, 2005, 17:29:58 ; Search time 3425 Seconds  
(without alignments)  
3678.352 Million cell updates/sec

Title: US-10-780-002-2  
Perfect score: 1368  
Sequence: 1 MKEGKHFLVHGACGGWSW.....MLCEPQKLCASLLEIAHKYN 260

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delgap 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/cgn2\_1/USFTO.spool\_p/US10780002/runat\_11092005\_104747\_26597/app.query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002\_@CIGN\_1\_1\_5600 @runat\_11092005\_104747\_26597 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.hcg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1368	100.0	1079	8 AY485932 Nicotiana
2	811	59.3	961	8 AF178576 Rauwolfia
3	793	58.0	1077	8 AF269158 Citrus si
4	768	56.1	792	6 AX506625 Sequence

5	768	56.1	792	8	BT014881	Arabidops
6	767	56.1	935	8	BT002859	Arabidops
7	734	53.7	891	8	AY751530	Catharant
8	696	50.9	870	8	AY086590	Arabidops
9	695	50.8	792	6	AX506856	Sequence
10	695	50.8	792	8	AY056068	Arabidops
11	695	50.8	932	8	AF361627	Arabidops
12	690	50.4	792	6	CQ806104	Sequence
13	690	50.4	792	6	AX505522	Sequence
14	690	50.4	792	8	BT015031	Arabidops
15	690	50.4	878	8	BT014901	Arabidops
16	680.5	49.7	792	8	BT010429	Arabidops
17	680.5	49.7	902	8	AK175772	Arabidops
18	669.5	48.9	1073	8	AK061058	Oryza sat
19	661.5	48.4	924	8	BT004125	Arabidops
20	645	47.1	783	8	BT012414	Arabidops
21	645	47.1	944	8	BT010709	Arabidops
22	640.5	46.8	770	6	AX651280	Sequence
23	640.5	46.8	771	6	AX506583	Sequence
24	640.5	46.8	771	6	BT006227	Arabidops
25	640.5	46.8	928	8	AK117107	Arabidops
26	635.5	46.5	123360	8	AC003040	Arabidops
27	635	46.4	777	8	AY093714	Arabidops
28	635	46.4	808	8	AY142490	Arabidops
29	634	46.3	928	8	AY046015	Arabidops
30	632.5	46.2	795	6	AX654444	Sequence
31	629	46.0	911	8	AY058115	Arabidops
32	605	44.2	789	8	AY455313	Lycopersi
33	590	43.1	828	8	AY142031	Arabidops
34	590	43.1	904	8	AY074858	Arabidops
35	589	43.1	903	8	AY086981	Arabidops
36	585	42.8	1078	8	HB040402	Hevea bras
37	585	42.8	1091	6	AS9586	Sequence 1
38	574.5	42.0	978	6	AX653408	Sequence
39	571.5	41.8	198780	8	ATC8RIV86	Arabidops
40	571.5	41.8	206420	8	ATAP21	Arabidops
41	564	41.2	57180	8	ATF18D22	Arabidops
42	554.5	40.5	1087	8	BT012867	Lycopersi
43	548	40.1	860	6	AX654758	Sequence
44	548	40.1	948	8	OSPIR7BR	Arabidops
45	540.5	39.5	142239	8	AF0033266	Oryza sat

ALIGNMENTS

RESULT 1	AY485932	Nicotiana tabacum salicylic acid-binding protein 2 mRNA	linear	PLN 07-JAN-2004
LOCUS	AY485932	1079 bp	mRNA	
DEFINITION	AY485932	Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.		
ACCESSION	AY485932			
VERSION	AY485932.1	GI:40549302		
KEYWORDS				
SOURCE		Nicotiana tabacum (common tobacco)		
ORGANISM		Nicotiana tabacum		
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana.		
AUTHORS		Kumar,D. and Klessig,D.F.		
TITLE		High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity		
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)		
PUBMED		14673096		
REFERENCE		2 (bases 1 to 1079)		
AUTHORS		Kumar,D. and Klessig,D.F.		
TITLE		Direct Submission		
JOURNAL		Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA		
FEATURES		Location/Qualifiers		
		1..1079		
		/organism="Nicotiana tabacum"		
		/mol_type="mRNA"		



```
Df 156 GCCGGCATCAACCAAGAGCTCGATGAGATTACACATTTGGGGACTACTCGGAGCCC 215
Qy 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Df 216 TTGATGGAAGTCATGCTAGTATCTCTCTGATGAGAGGTGTCTTCTTGCCATAGC 275
Qy 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Df 276 TTTGGTGGCATGAGTTGGTCTTGCCATGGAAACCTACCCAGAGAAGATATCAGTTGCT 335
Qy 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
Df 336 GTTTTATGCTGCAATGATGCTGATCTCTTAACCACTCACTATCTATCGTTTGAGAAG 395
Qy 122 TyrAsnGluArgThrProAlaGluAsnTyrPheLeuAspThrGlnPheLeuProTyrGlySer 141
Df 396 TACATGAGAAGTGTCCGACGATATGATGTTGGACTCACAGTTTCAACCTACCGAAGC 455
Qy 142 ProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyr 161
Df 456 CCAGAGAACCCAGGAATGTCATGATTCTTGGACCTCAGTTTATGGCCCTCAAAATGTTTC 515
Qy 162 GlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeu 181
Df 516 CAGAAATTGCTCAGTCAGGACCTTGAATTTAGCCAAATGTTGACTCGACAGGTTGCTTA 575
Qy 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201
Df 576 TTTTCCAGATTGTCAGGAGGCCAAAGTTCTCAACCGAGAGTACGTTCCGGTGAAG 635
Qy 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGln 221
Df 636 CGAGCTTATATCTTTTGCATGAAGATAAATCATTTCCAGTTGAGTTTCAGAAATGGTTT 695
Qy 222 IleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaHisMetAlaMet 241
Df 696 GTTGAAGTGTGGAGCTGATTAAGTAAAGAAAGAAATCAAGAGACGAGATCATATGGGAATG 755
Qy 242 LeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAla 256
Df 756 CTTTCGACGCAAGGAAGTTTGCAAAGTCCTGCTGATATATCA 800
```

```
RESULT 3
LOCUS AF269158
DEFINITION Citrus sinensis ethylene-induced esterase mRNA, complete cds.
ACCESSION AF269158
VERSION AF269158.1 GI:14279436
KEYWORDS .
SOURCE Citrus sinensis
ORGANISM Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCES 1 (bases 1 to 1077)
AUTHORS Zhong,G.Y., Goren,R., Riov,J., Sisler,E.C. and Holland,D.
TITLE Characterization of an ethylene-induced esterase gene isolated from
Citrus sinensis by competitive hybridization
JOURNAL Physiol. Plantarum 113 (2), 267-274 (2001)
PUBMED 12060305
REFERENCE 2 (bases 1 to 1077)
AUTHORS Zhong,G.Y., Goren,R., Riov,J. and Holland,D.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-2000) Horticulture, Faculty of Agricultural, Food
and Environmental Quality Sciences, The Hebrew University of
Jerusalem, Rehovot 76100, Israel
FEATURES
Location/Qualifiers
1..1077
/organism="Citrus sinensis"
/mol_type="mRNA"
/db_xref="taxon:2711"
55..858
/notes="EIE"
```

```
/codon_start=1
/product="ethylene-induced esterase"
/protein_id="AAKS8599.1"
/db_xref="GI:14279437"
/translation="MEEVVGMEEKHFVLVHGNNHCANWCYKLKARLVAGGHRVTAVDL
AASGNNMRIEDVHTFHAYSPLEMEVLASLPABEKVILVHSLGGVTIALAGKPEPHK
ISVAVFTAFNPDTTHRPSFVLEQYSEKMGKDDSWLDTQFSQCDASPHISMFLPGR
EFLTKIYQLCPDPDLKALMLVPGSMFIDNLSKESFSGDEGYSKRVYVLVCEEDI
GLPKQFQHMWIIQNYPVNEVMEIKGSDHMAMLSDPKQLCDCLSLQISLKA"
```

Alignment Scores:	6e-64	Length:	1077
Pred. No.:	793.00	Matches:	152
Score:	73.36%	Conservative:	38
Percent Similarity:	58.69%	Mismatches:	67
Best Local Similarity:	57.97%	Indels:	2
Query Match:	8	Gaps:	1
DB:			

US-10-780-002-2 (1-260) x AF269158 (1-1077)

```
Qy 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyrLys 22
Df 76 GAAGAAGAGCATTTTGTCTAGTTTCATGGAGTAAACCATGGAGCATGGTTGGTACAAA 135
Qy 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaLaser 42
Df 136 CTGAAGCAAGCGCTGGTGGGGGGGTCCCGGGTGACGGCTGTGGACCTAGCCGCTCG 195
Qy 43 GlyThrAspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeu 62
Df 196 GGCATCAACATCAAGAGAAATTGAGGATGTGCACACATTCATGCATACAGTAGGCCCTTG 255
Qy 63 MetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeu 82
Df 256 ATGGAGGTTTGGGCATCATTCTCTCCGGAAGAAAGGTGCATCTTGTGGACACAGTCTT 315
Qy 83 GlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaVal 102
Df 316 GGGGGCGTCACTTTGGCCCTTGGCCGCGACAAATTCACACAAATCTCCGTGGCTGTT 375
Qy 103 PheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyr 122
Df 376 TCGTAACCTGCATTTCCTCTGACACACACACCGGCCATCTTTTGTGTTTGGAGCATAT 435
Qy 123 AsnGluArgThrProAlaGlu-----AsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Df 436 TCTGAGAAGATGGGAAAGAGGACGACAGCTGGTTGGACACTCAATTTTCACAATGTGAC 495
Qy 141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160
Df 496 CGGTCAATCCATCTCACATTTCCATGCTTTCCGCGCGAGTCTTTCAGCTATCAAGATC 555
Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Df 556 TATCAGCTTTGTCTCTCTGAGGATCTGAGGCTGGCAAGATGTTGGTGAGGCCAGATCA 615
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Df 616 ATGTTTATAGCAACTTATCGAAGGAAAGTAAAGTTTCAGCGATGAAGGATACCGATCTGTT 675
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Df 676 AAGCAGGTTTATCTGTATGCGAAGAGGATATTGCTCTCCCTAAGCAATTTTCAGCAGCTGG 735
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Df 736 ATGATCCAAAACATATCCCGTTAAATGAAGTGTGAGATCAAGGGCGGTGATCACATGGCA 795
Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Df 796 ATGCTTTCGACCCACAGAAACTTTTGGGATTTGTCTGTCTCAGATTTCTCTTAAGTAT 852
```

```
RESULT 4
AX506625
LOCUS AX506625 792 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1320 from Patent WO216655.
ACCESSION AX506625
VERSION AX506625.1 GI:23387862
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
TITLES Stems-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1320 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1. .792
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 8,986-62 Length: 792
Score: 768.00 Matches: 149
Percent Similarity: 71.04% Conservative: 35
Best Local Similarity: 57.53% Mismatches: 73
Query Match: 56.14% Indels: 2
DB: 6 Gaps: 2
US-10-780-002-2 (1-260) x AX506625 (1-792)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 13 AAGAGGAAACAACATTTTGTACTAGTACATGTTCTGTCGCCGCGTGGTGGTAC 72
QY 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThAlaLeuAspLeuAla 41
Db 73 AAGTTTAAGCCGCTGCTAGAGCGGTGGGCCACCGCGTAACTGCTGTGGACCTAGCTGCC 132
QY 42 SerGlyThrAsp---LeuArgLysLleGluGluLeuArgThrLeuTyAspTyrThrLeu 60
Db 133 TCCGGAATAGACACACAGAGGTGCTGATCTGATGATCCCCACATCGAACCAATCTCGGAG 192
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValLleLeuValGlyHis 80
Db 193 CCATTGACGAAGCTCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTGCTGGTCCAC 252
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysLleTyrAla 100
Db 253 ACCTTTGGTGGCTTGAACCTTACCATTAGCCATGACATGAAAGTTTCCGAAAATCTCTGTC 312
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 313 GCTGTATTCTTGACTGCTTTTCATGCGCGGACACCGAACCTCACCATCTCTCTTGGAC 372
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 373 AAGTTTGGGAAGCAACATGCTCAAGAAAGCATGATGGGACCGCAATTCGAACCTTATGCT 432
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 433 TCAGACAATTCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 489
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 490 TACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGGCCAGATCG 549
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
```

```
Db 550 TTATTTATTAACGATTATTCGAAGATGAAACATCTTCGGATGAAGGATATGGGTCTGTT 609
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 610 CCTCGAGTTTTCATAGTGTGTAAGAGGACAAACCAATTCACAGAAACGCCAGAGATGG 669
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240
Db 670 ATGATTGATAATTTTCGGTGAATTTAGTGTGATGGATGGAGGACAGATCATATGCCA 729
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 730 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAATCGCGGACAAATTC 786
RESULT 5
LOCUS BT014881 792 bp mRNA linear PLN 05-JUN-2004
DEFINITION Arabidopsis thaliana At2g23620 gene, complete cds.
ACCESSION BT014881
VERSION BT014881.1 GI:48310670
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 792)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 792)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
Direct Submission
Submitted (05-JUN-2004) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
FEATURES
source
1. .792
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="2"
/ecotype="Columbia"
/clone="U67545"
/notes="This clone is in pUNI 51"
1. .792
/notes="unknown protein"
/codon_start=1
/product="At2g23620"
/protein_id="AA041864.1"
/db_xref="GI:48310671"
/translation="MSSEKRRKQHFVLVHSGCHGACWYKVKPLLEAVGHVRVTAVDLAA
SGITPTSIDTPTCEQYSEPLTKLLTSLPNDEKVLVGHSGFGLNLAIAMEKRPKEI
SVAVPLTAFMEDTEHSFSLVLDKFGSNWPQAWMGTEPEPYGSDNSGLSMFSPDFMK
LGLYQLSPDELDLEGLLMLRPSLFIINDLSKKKNFSDGEGYSVPRFVICKEDKAIFE
ERQRWMDINDFPNVLVMEEMETDHMFCKPOOLSDFLUKIADKFV"
ORIGIN
Alignment Scores:
Pred. No.: 8,986-62 Length: 792
Score: 768.00 Matches: 149
Percent Similarity: 71.04% Conservative: 35
Best Local Similarity: 57.53% Mismatches: 73
Query Match: 56.14% Indels: 2
DB: 8 Gaps: 2
US-10-780-002-2 (1-260) x BT014881 (1-792)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 13 AAGAGGAAACAACATTTTGTACTAGTACATGTTCTGTCGCCGCGTGGTGGTAC 72
```



```
US-10-780-002-2 (1-260) x BT002859 (1-935)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 56 AAGAGGAACAACATTTTGTACTAGTACATGGTTCGTCCATGGCGCGTGGTCTGGTAC 115
QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 116 AAGGTTAAGCCGCTGTAGAGCGGTGGCCACCGCGTAACCTGCTGTGGACTAGCTGCC 175
QY 42 SerGlyThrAsp--LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 176 TCCGGAATAGACAACAGAGGTGATCTACTGACATCCCACATGCGAACAATACTCGGAG 235
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValLileLeuValGlyHis 80
Db 236 CCATTGACGAAGCTCTGACCTCATTCGCAAAATGATGAAGAGTTGTGCTGGTGCAC 295
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 296 AGCTTTGGTGGCTTGAACCTAGCCATAGCCATGGAAAAGTTTCCCAAAAAATCTCTGTC 355
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 356 GCTGTATTCTTGACTGCTTTCATCGCGACACCGAACACTCACCATCTCTGCTTGGAC 415
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 416 AAGTTTGGNAGCAACATGCTCAAGNAGCATGGATGGCGACCGAATTCGAACCTTATGTT 475
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 476 TCAGACAATTCGGGACTG---AGTATGTTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 532
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 533 TACCAGCTTCTCCAGTTGAGGATCTTGAACTGGGATTACTTTTATGAGGCCAGATCG 592
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 593 TTATTATTATACGATTTATCGAAGATGAAAACCTCTCGATGAAGGATATGGGTCTGTT 652
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 653 CCTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATTCACAGAAAGCCGACAGATGG 712
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 713 ATGATTGATAAATTTCCGGTGAATTTAGTGTGGAGATGGAGGACAGATCATATGCCA 772
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 773 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTTGAAATTCGGGACAAATTC 829

RESULT 7
AY751530 891 bp mRNA linear PLN 11-OCT-2004
LOCUS
Catharanthus roseus protein S mRNA, complete cds.
DEFINITION
AY751530
ACCESSION
VERSION
AY751530.1 GI:53830669
KEYWORDS
Catharanthus roseus (Madagascar periwinkle)
SOURCE
Catharanthus roseus
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vinceae; Catharanthus.
REFERENCE
1 (bases 1 to 891)
Lemenager, D., Rideau, M. and Clastre, M.
AUTHORS
Cloning of the Catharanthus roseus protein S (CrPS) associated with
TITLE
monoterpene indole alkaloid production
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 891)
Lemenager, D., Rideau, M. and Clastre, M.
AUTHORS
Direct Submission
TITLE
Submitted (15-SEP-2004) Laboratoire de Biologie Molculaire et
JOURNAL
Biochimie Vegetale, Faculte de Pharmacie, 31 Avenue Monge, Tours
37200, France
FEATURES
Location/Qualifiers
source
1..891
/organism="Catharanthus roseus"
/mol_type="mRNA"
/db_xref="taxon:4058"
CDS
48..824
/note="hydrolase; CrPS; alpha/beta fold family protein"
/codon_start=1
/product="protein S"
/protein_id="AAU5203.1"
/db_xref="GI:53830670"
/translation="MEYKHFVTVHGVGHGAVVYKLPRIEAAAGHRCTAVNLAAGI
NEKLSVRISIDYAAPLEVLDSVPENKIVLVHSGGMMTAAVGMKFPNKISLAV
FLNALPDTNRNPSYVLEEVYAKTPPEAWKDCQFSAYGDPPIITSLVCGPEFISSTLVH
LSPIEDHALGKILVRPSLFTEDLLKAEKTEEGFSGVPRVYVIAEDKTIPTPEFQW
MIENNPVKEVKEIKGADHMPMPFKPDELSQLLDIAKHA"
```





SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1  
 REFERENCE Harper, J.F., Kress, J., Wang, X. and Zhu, T.  
 AUTHORS Stress-regulated genes of plants, transgenic plants containing  
 TITLE same, and methods of use  
 JOURNAL Patent: WO 0216655-A 1551 28-FEB-2002;  
 The Scripps Research Institute (US); Syngenta Participations AG  
 (CH)  
 FEATURES Location/Qualifiers  
 source 1..792  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"  
 ORIGIN  
 Alignment Scores: Length: 792  
 Pred. No.: 695.00 Matches: 138  
 Score: 67.18% Conservative: 36  
 Percent Similarity: 53.28% Mismatches: 83  
 Best Local Similarity: 50.80% Indels: 2  
 Query Match: 6 Gaps: 2  
 DB: 2  
 US-10-780-002-2 (1-260) x AX506856 (1-792)  
 QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21  
 DB 13 AAGAGAAGCAACACTTCGTGTAGTACATGGTGGCCGCGCATGGTCTGGTAC 72  
 QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
 DB 73 AAGTTAAGCCCTCTTCGAGGCTTTGGSCCATCGTGAACCGCTTAGACCTAGCTGCT 132  
 QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuLysAspTyrThrLeu 60  
 DB 133 TCCGGTATAGACACACACAGGTCATCTACATCTTCTACATGTGCAACATATCTGAG 192  
 QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
 DB 193 CATTGTATGCACCTAATGACTTTCATTCGCGAATGATGAGAGGTTGTACTCTGTTGTCAT 252  
 QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 DB 253 AGCTTTGGAGGTTTGAGTTTAGCTTAGCCATGGATAAGTTCCCGATAAAATCTCTGTC 312  
 QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGlu 120  
 DB 313 TCTGTCTTCGACTGCATTCATGCCGACACCAACACTCACCATCGTTCTCGAGGAA 372  
 QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 DB 373 AAGTTTGCACGACGATGACACACAGAGGATGGATGGCTCTGAGCTCGAGACATATGTT 432  
 QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPhePheLeuAlaHisLysLeu 160  
 DB 433 TCAGATAATTCGGCTTG--TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCGTCTC 489  
 QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
 DB 490 TACCAACTTTCCTCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCGCTAGTTCA 549  
 QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 DB 550 TTGTTTATTAAAGATTATCGAAGATGGAGAACTTTCTGAGAAAGGCTATGACTGTT 609  
 QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 DB 610 CCTCGAGCTTACATTGTGTGCAAGAGGACCAACATTATCTCGAGAGACCATCAACGATGG 669

QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
 DB 670 ATGATCATATATTTATCCGCGAATTTAGTATTGAGATGGAGAGACTGATCATATGCCA 729  
 QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 DB 730 ATGTTTGGCAACCTCAACTACTAAGTGACCATCTATTGGCAATCGTGCAATTTTC 786  
 RESULT 10  
 AY056068 792 bp mRNA linear PLN 30-SEP-2001  
 LOCUS Arabidopsis thaliana At2g23600/F26B6.25 mRNA, complete cds.  
 DEFINITION AY056068  
 ACCESSION AY056068.1 GI:15810084  
 VERSION FLI\_CDNA.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Arabidopsids thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 792)  
 Koesema, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P.,  
 Banth, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,  
 Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A.,  
 Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,  
 Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,  
 Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,  
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
 Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsids ORF clones  
 Unpublished  
 2 (bases 1 to 792)  
 Koesema, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P.,  
 Banth, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,  
 Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A.,  
 Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,  
 Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,  
 Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,  
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
 Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.  
 Direct Submission  
 Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory  
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN  
 Arabidopsis Full-Length cDNA) : Seki, M., Narusaka, M., Ishida, J.,  
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
 Hayashizaki, Y. and Shinzaki, K.  
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Koesema, E., Chen, H.,  
 Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P., Banth, J., Bowser, L.,  
 Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,  
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
 Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,  
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
 Davis, R.W., Theologis, A., and Ecker, J.R.  
 Koesema, E. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally  
 to this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
 contributed equally to this work as PIs.  
 FEATURES Location/Qualifiers  
 source 1..792  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="2"  
 /clone="U12118"  
 /ecotype="Columbia"  
 /note="This clone is in pUNI 51"  
 1..792

CDS

/notes="putative acetone-cyanohydrin lyase"

/codon\_start=1  
/product="At2g23600/F26B6.25"  
/protein\_id="AAL06968.1"  
/db\_xref="GI:15810085"  
/translation="MSEERKKQHFLVHGACGAWCYKPKLLEALGHRVTALDLAA  
SGIDTTRSTIDTSCQYSEPLQMLTSLPNDEKVLVGHSGFGLSLALAMDKFPDKI  
SVSVFTAFMPDTHKSPFSFVEEKFASSMTPEGMWSELETYGDNSGLSVFSTDFMK  
HRLYQLSPVEDLGLLLKRPSSLFINELSKMENPEKGYGVPVPRAYIVCKEDNIIE  
DQRMWHNYPANLVIMBETDHMPCKPQLLSHLLAIADNFC"

ORIGIN

Alignment Scores:			
Pred. No.:	5.08e-55	Length:	792
Score:	695.00	Matches:	138
Percent Similarity:	67.18%	Conservative:	36
Best Local Similarity:	53.28%	Mismatches:	83
Query Match:	50.80%	Indels:	2
DB:	8	Gaps:	2
US-10-780-002-2 (1-260) x AY056068 (1-792)			
Qy	2	LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr	21
Db	13	AAGAGGAAGCAACACTTCGTGCTAGTACATGGTGCACCGCGCATGGTCTGTGATC	72
Qy	22	LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla	41
Db	73	AAGTTAAGCCTCTTCTCGAGGCTTTCGGCCATCGTGTACCGCCCTAGACCTAGCTGCT	132
Qy	42	SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu	60
Db	133	TCCGGTATAGACACACAGGTCAATCACTGACATTCCTACATGTGCAACAATTCCTGAG	192
Qy	61	ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis	80
Db	193	CCATTGATCGAGCTAAATGACTTCATTCGCCGAATGATGAGAAGGTGTACTCGTGTGTCAT	252
Qy	81	SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla	100
Db	253	AGCTTGGAGGTTTGAGTTTGGCTTACCATGGATGGATAGTTTCCCGATAAATCTCTGTC	312
Qy	101	AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu	120
Db	313	TCTGTCTCTCGTGCATGCTATTCATGCCCGACACCAACACTCACCATCGTTCGTCGAGGA	372
Qy	121	GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly	140
Db	373	AAGTTTGCACAGCAGCATGACACACCAAGGATGATGGCTCTGAGCTCGAGACATATGCT	432
Qy	141	SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu	160
Db	433	TCAGATAATTCCGGCTTG---TCTGTGTTCTTCAGCACCAGCTTCATGAAGCACCGCTC	489
Qy	161	TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer	180
Db	490	TACCAACTTCTCTCTGGAGGATCTTGACCTTGGATTGCTCTAAAGAGCGCTAGTTCA	549
Qy	181	LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal	200
Db	550	TTGTTTTTAATGAATTCATGAGATCGGAACTTTTTCGAGAAAGGATGATGATCTGTT	609
Qy	201	LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp	220
Db	610	CCTCGAGCTTACATTTGTGCAAGAGGACAACTTATCTCGGAAGACCATCAACGATGG	669
Qy	221	GlnIleAspAsnIleGlyValThrGluAlaIleLeuIleLysGlyAlaAspHisMetAla	240
Db	670	ATGATCCATAATTCATCCGCGGAATTTAGTCATTGATGATGAGACGACTGATCATATGCCA	729
Qy	241	MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr	259
Db	730	ATGTTTTGCAAACTCAACTACTTAAGTGACCATCTATTGGCAATCGCTGACAATTTC	786

RESULT 11  
AF361627  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF361627 932 bp mRNA linear PLN 07-FEB-2002  
Arabidopsis thaliana At2g23600/F26B6.25 mRNA, complete cds.  
AF361627  
FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE  
AUTHORS

1 (bases 1 to 932)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,  
Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M.,  
Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinozaki, K.,  
Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 932)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,  
Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M.,  
Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinozaki, K.,  
Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE  
JOURNAL

Submitted (15-MAR-2001) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,  
Cheuk, R., Kim, C.J., Koseema, E., Meyers, M.C., Tracy, S.E., Banh, J.,  
Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,  
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,  
Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,  
and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
contributed equally to this work as PIs.

FEATURES  
source

1..932  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAFL09-12-L06 (R12118)"  
/ecotype="Columbia"  
1..13  
14..805  
/note="putative acetone-cyanohydrin lyase"  
/codon\_start=1  
/product="At2g23600/F26B6.25"  
/protein\_id="AAK32795.1"  
/db\_xref="GI:13605603"  
/translation="MSEERKKQHFLVHGACGAWCYKPKLLEALGHRVTALDLAA  
SGIDTTRSTIDTSCQYSEPLQMLTSLPNDEKVLVGHSGFGLSLALAMDKFPDKI  
SVSVFTAFMPDTHKSPFSFVEEKFASSMTPEGMWSELETYGDNSGLSVFSTDFMK  
HRLYQLSPVEDLGLLLKRPSSLFINELSKMENPEKGYGVPVPRAYIVCKEDNIIE

5'UTR  
CDS

```
3' UTR
ORIGIN
DHORWMIHNPANLVIEEMETDMPFCKPQLLSHLLAIADNFC"
806..932

Alignment Scores:
Pred. No.: 6,01e-55 Length: 932
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x AF361627 (1-932)

QY 2 LysGluGlyLeuPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 26 NAGAGGAGCAACACTCTCGTGTAGTACATGCTGCGTCCAGCGCATGGTGTGTAC 85

QY 22 LysLeuLeuProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 86 AAGGTTAAGCCTCTTCGAGGCTTTGGGCCATCGTGTAAACCGCCTTAGACCTAGCTGCT 145

QY 42 SerGlyThrAsp--LeuArgLysIleGluLeuLeuArgThrLeuTyrAspTyrThrLeu 60
Db 146 TCGGTATAGACACACACAGGTCATCACTGACATTTCTACATGTGAACATATTTCTGAG 205

QY 61 ProLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 206 CCATTGATGACGTAATGACTTCAATTCGCGAATGATGAGAGGTTGTACTCGTTGGTCA 265

QY 81 SerLeuGlyGlyMetAsnLeuLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 266 AGCTTTGGAGGTTTGAGTTAGCCTTAGCCATGGATGAAGTTTCCCGATAAAATCTCTGC 325

QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 326 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385

QY 121 GlnTyrAsnGluArgThrProAlaGlnAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 386 AAGTTTGAAGCAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 445

QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 446 TCAGATAATTCGGCTTG--TCTGTGTTCTTCAGCACCAGCTTCATGAAGCACCCTCTC 502

QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 503 TACCAACTTCTCCTGTGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTAGTTCA 562

QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 563 TTGTTTATTAAATTAATTCGAAGATGGAGAACTTTCTCGAAGAGGATGGATGCTGTT 622

QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 623 CCTCGAGCTTACATTGTGCAAGAGGACCAATTTATCTCGGAAGACCATCAACGATGG 682

QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 683 ATGATCCATTAATATCCGCGGAATTAGTGTATGATGGAGAGAGCTGATCATATGCCA 742

QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 743 ATGTTTGCACAACTCACTACTACTAGTGACCATCTATTGGCAATCGCTGACAAATTC 799

RESULT 12
CQ806104
LOCUS
DEFINITION
Sequence 2515 from Patent WO2004035798.
ACCESSION
CQ806104
VERSION
CQ806104.1 GI:47111686
KEYWORDS
```

```
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
1
AUTHORS
Inze, D., de Veylder, L. and Vlieghe, K.
TITLE
Identification of novel e2f target genes and use thereof
JOURNAL
Patent: WO 2004035798-A 2515 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
Location/Qualifiers
1..792
source
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 1.47e-54 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x CQ806104 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
Db 22 CAACACGTCGTTCTAGTACATGCTTGCCTCCATGGCGCTGGTGTGTAAG 81

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44
Db 82 CCGAGCTCGAGGCTTTCGCCACCGGTAAACCGGTAGATCTAGCTCCCTCGGTATA 141

QY 45 AspLeu--ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 142 GACATGACCAGTCATACAGATATATCCATCGCAACATATCTCAGAGCCATTGATG 201

QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 202 CAGCTAATGACTCCTACCTACCATGATGATGAGAGGTTGTGCTTGTGTGCTAGCTTAGA 261

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db 262 GGTTTGATTTAGCTATGCGCATGATGATTTCCGACCAAAATCTGTTTCTGTCITT 321

QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123
Db 322 GTGACTGTATGATGCCAGACACCAACACTCACCATCTTCGTATGGGATAAGTAAGA 381

QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
Db 382 AAGAAACTTTCACGAGAGGAATGGTTAGACACCGTGTTCAGGACCTGAT 438

QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
Db 439 TTTCTCTACGAGTTTGGATTTTGGACCAATTCATGGCCCAAGAACTTGTATCAGTTG 498

QY 164 CysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
Db 499 TCTCCAGTCCAAGATCTTGAATTCGCGAAATTTGTTGGGGCAACCCATTGATTAAG 558

QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db 559 AAGATATGGCAGAGAGAGAGGTTTCTAGTGAGGAAGGATCCGGATCCGTTACAGCTATA 618

QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 619 TTTATTGTATCGGAAGAGGATCTTGTGTACCCCGAAGATTACACCGATCGATGATCAGC 678

QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
```

```
Db 679 AACTTTCCCCAAAGAAGTAATGGAGATCAAGACGACAGATCATATGCGAATGTTCTCC 738
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluLeuAlaHisLysTyr 259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 739 AAGCTCAACAACACTATGCTCTCTCTTGAGATTGCAATAATAT 786

RESULT 13
LOCUS AX505522 792 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 217 from Patent WO0216655.
ACCESSION AX505522
VERSION AX505522.1 GI:23386759
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 217 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source
    1..792
    /organism="Arabidopsis thaliana"
    /mol_type="unassigned DNA"
    /db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 1,478-54 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: 6 Gaps: 2

US-10-780-002-2 (1-260) x AX505522 (1-792)
QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrLysLeuLys 24
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 22 CAACAGCTGCTTCTAGTACATGCTGTGCCATGGCGCTGCTGGTACAGGTAG 81

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 82 CCGCAGCTCGAGGCTTCTGGCCACCGCTAACCGCGTAGATCTAGCTGCTCCGGTATA 141

QY 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 142 GACATGACCAAGGTCATACAGATATATCCATATATCCATGCGAACCAATCTCAGAGCCATTGATG 201

QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 202 CAGCTAATGACCTCACTACAGATGATGAGAGAGTTGCTGTTGGTCTAGCTAGGA 261

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 262 GGTTTCAGTTTACGTATGCCATGGATATGTTTCCGACCAAAATCTCTGTTCTGCTTT 321

QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 322 GTGACTGTCTATGATGTCAGACACCAACCACTCACCATCTCTGATGGATAGCTAAGA 381

QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 382 AAGAAACTTCACGAGAGGAATGTTAGACCGTGTTTACAGC---GAGAAACCTGAT 438

QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 439 TTTCCTAGCAGGTTTGGATTGTCGACCAAGATTATCGCCCAAGAACTTGTATCAGTTG 498
```

```
QY 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 499 TCTCCAGTCAGAGATCTGAATGGCGAAAATGTTGGTGGCGCAACCCATGATTAAG 558

QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 559 AAGATATATGCAGAGAGAAGAGCTTCAGTAGGAAGGATACGATCCGTTACACGTATA 618

QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 619 TTTATTGTATGCGAAAGGATCTTGTGTCCACCGAAGATTACCAGCGATCGATGATCAGC 678

QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 679 AACTTTCCCCAAAGAAGTAATGGAGATCAAGACGACAGATCATATGCGAATGTTCTCC 738

QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 739 AAGCCTCAACAACACTATGCTCTCTCTTGGAGATTGCAATAATAT 786

RESULT 14
LOCUS BT015031 792 bp mRNA linear PLN 10-JUL-2004
DEFINITION Arabidopsis thaliana At2g23610 mRNA, complete cds.
ACCESSION BT015031
VERSION BT015031.1 GI:50198957
KEYWORDS FLI-CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 792)
REFERENCE
AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE
AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2004) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
FEATURES
source
    1..792
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"
    /db_xref="taxon:3702"
    /chromosome="2"
    /clone="U67063"
    /ecotype="Columbia"
    /note="This clone is in pUNI 51"
CDS
    1..792
    /note="unknown protein"
    /codon_start=1
    /product="At2g23610"
    /protein_id="AA70482.1"
    /db_xref="GI:50198958"
    /translation="MSEERKQHVVLVHGACHGAWCYKVKPQLEASGHRVTVADLAA
SGIDWTRISITDITCEQYSEPLMOLMTSLPDKEKVLVGHISGGLSLAMAMDMPFTKI
SVSVFTAMDPDTHKSPSPFWDKLRKETSREWLDTVETSKPDPSPSEFWIPGPPEMA
KNLYQLSPQDLELAKMLVRANPLIKDMABRRSFSEEGYGSVTRIFIVCGKDLVSPE
DYQRSMISNFPFPEKMEIKDADHMPFMSKPOLCALLLEIAANKYA"
ORIGIN
Alignment Scores:
Pred. No.: 1,478-54 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: 8 Gaps: 2
```

US-10-780-002-2 (1-260) x BT015031 (1-792)

```
Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrPserTyrTyrLysLeuLys 24
Db 22 CAACACGTCGTTCTAGTACATGGTCTTGCCATGGCGCTGGTGTGGTACAGGTTAAG 81
Qy 25 ProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaAAserGlyThr 44
Db 82 CGCAGCTCGAGGCTTCTGGCCACCGCTAACCGCGTAGATCTAGCTGCCTCCGCTATA 141
Qy 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 142 GACATGACGACGATCAATCAGATATATCCATCCGAAACAATCTCAGAGCATTTGATG 201
Qy 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 202 CAGCTTAATGACCTCACTACCAGATGATGAGAAGGTTGCTTGTGGTCAATAGCTTAGGA 261
Qy 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db 262 GGTGTTGAGTTAGTATGATGCAATGATATGTTCCGACCAAAATCTCTGTTCTGCTTT 321
Qy 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
Db 322 GTGACTGCTATGATCCAGACACCAACACTCACCATCCTTCGTATGGATGAAGCTAAGA 381
Qy 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
Db 382 AAAGAAACTTCCAGAGAGAAATGGTTAGACACCGTGTATTACGAGC---GAGAAACCTGAT 438
Qy 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
Db 439 TTTCTCTAGGAGCTTTTGGATTTTGGACCAAGAATTCATGCCCAAGAATCTGTATCAGTTG 498
Qy 164 CysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
Db 499 TCTCCAGTCACAGATCTTGAATGGCGAAATGTTGGTGGAGGCAAAACCATTTGATTAG 558
Qy 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db 559 AAAGATATGGCAGAGAGAGAGCTTCAGTGAGGAGGATACGGATCCGTTACACGTATA 618
Qy 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 619 TTTATTTGATGCGGAAAGATCTTGTGTCAACCGAAGATTACACGCGATCGATGATCAGC 678
Qy 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 679 AACTTTCCCCAAAAGAGTAATGGAGATCAAGACGCGAGATCATATGCCAATGTTCTCC 738
Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 739 AAGCCTCAACACTATGTGCTCTTCTCTTGGAGATTGCAATAAATAT 786
```

## RESULT 15

BT014901 878 bp mRNA linear PLN 11-JUN-2004  
LOCUS Arabidopsis thaliana At2g23610 gene, complete cds.

## ACCESSION

BT014901

## VERSION

BT014901.1

## KEYWORDS

FLI CDNA.

## SOURCE

Arabidopsis thaliana

## ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 878)

Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 878)

Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

```

Qy 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db 566 AAAGATATGCGAGAGAGAGAGCTTCAGTCAGGAGGATACGGATCCGTTACACGTATA 625
Qy 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 626 TTTATTGTATGCGGAAAGGATCTTGTGTACCCGAGATTACCAGCGATCGATCATCAGC 685
Qy 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 686 AACTTCCCCCAAGAGAGTAATGGAGATCAAGACGAGATCATATGCCAATGTTCTTCC 745
Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 746 AAGCCTCAACAACATATGTGCTTCTCTTGGAGATTGCAATAAATAT 793

```

Search completed: September 12, 2005, 21:55:34  
Job time : 3435 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 15:37:37 ; Search time 733.243 Seconds  
(without alignments)  
9663.988 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaagaaaa.....gaaaaaaaaaaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 7351250 seqs, 3283620254 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1079	100.0	1079	21	US-10-780-002-1
2	29	2.7	607	13	Sequence 1, Appli
3	29	2.7	607	17	Sequence 241778,
4	28	2.6	615	20	Sequence 241778
5	27	2.5	207	17	Sequence 79873, A
6	27	2.5	207	18	Sequence 17538, A
7	27	2.5	321	9	Sequence 17538, A
					Sequence 4280, Ap

C	8	27	2.5	446	10	US-09-873-367C-946	Sequence 946, App
C	9	27	2.5	446	21	US-10-843-641A-946	Sequence 946, App
C	10	27	2.5	511	20	US-10-425-115-102591	Sequence 102591,
	11	27	2.5	1795	9	US-09-925-300-754	Sequence 754, App
	12	27	2.5	1795	17	US-10-264-237-1253	Sequence 1253, App
	13	26	2.4	332	17	US-10-242-535A-54828	Sequence 54828, A
	14	26	2.4	332	18	US-10-085-783A-54828	Sequence 54828, A
C	15	26	2.4	379	19	US-10-437-963-72087	Sequence 72087, A
C	16	26	2.4	406	20	US-10-425-115-111034	Sequence 111034,
	17	26	2.4	814	9	US-09-780-717-34	Sequence 34, Appl
	18	26	2.4	1099	22	US-10-487-092-68	Sequence 68, Appl
	19	26	2.4	98345	21	US-10-461-862-136	Sequence 136, App
	20	25	2.3	135	17	US-10-242-535A-3360	Sequence 3360, App
	21	25	2.3	135	18	US-10-085-783A-3360	Sequence 3360, App
C	22	25	2.3	148	20	US-10-357-930-60780	Sequence 60780, A
	23	25	2.3	157	17	US-10-242-535A-17475	Sequence 17475, A
	24	25	2.3	157	18	US-10-085-783A-17475	Sequence 17475, A
C	25	25	2.3	158	20	US-10-357-930-61029	Sequence 61029, A
C	26	25	2.3	172	20	US-10-425-115-166206	Sequence 166206,
	27	25	2.3	182	20	US-10-425-115-85062	Sequence 85062, A
C	28	25	2.3	194	20	US-10-357-930-60898	Sequence 60898, A
C	29	25	2.3	214	9	US-09-983-965-591	Sequence 591, App
	30	25	2.3	219	9	US-09-764-846-51	Sequence 51, Appl
	31	25	2.3	219	14	US-10-091-483-51	Sequence 51, Appl
C	32	25	2.3	221	20	US-10-357-930-61823	Sequence 61823, A
C	33	25	2.3	257	20	US-10-425-115-155173	Sequence 155173,
C	34	25	2.3	258	18	US-10-424-599-123984	Sequence 123984,
	35	25	2.3	276	20	US-10-425-115-114837	Sequence 114837,
	36	25	2.3	282	14	US-10-066-543-2060	Sequence 2060, Ap
C	37	25	2.3	290	20	US-10-357-930-61438	Sequence 61438, A
C	38	25	2.3	298	19	US-10-437-963-90281	Sequence 90281, A
C	39	25	2.3	310	20	US-10-723-860-1322	Sequence 1322, Ap
C	40	25	2.3	319	18	US-10-424-599-73498	Sequence 73498, A
C	41	25	2.3	324	20	US-10-425-115-135371	Sequence 135371,
C	42	25	2.3	341	10	US-09-814-353-4668	Sequence 4668, Ap
C	43	25	2.3	341	10	US-09-814-353-10967	Sequence 10967, A
C	44	25	2.3	341	20	US-10-425-115-142143	Sequence 142143,
C	45	25	2.3	356	18	US-10-424-599-77760	Sequence 77760, A

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US/107800002  
; Publication No. US2005034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE OF INVENTION: ENCODING NUCLEIC ACIDS, SABP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-10-780-002-1

Query Match 100.0%; Score 1079; DB 21; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGCCGGGAAAGAAAGAAACTTAAACAGGCATAAATTCAAATGAAGGAAACACT 60  
|||||

Db 1 ACGCGGGAAGAAAGAACTAACAGGCATATAAATTCAAATGAAGNAGGAAACACT 60  
Qy 61 TTGTTTTAGTACATGGTGCATCCCATGGAGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
Db 61 TTGTTTTAGTACATGGTGCATCCCATGGAGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
Qy 121 TAGAAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTCTGGCACTGATTTGA 180  
Db 121 TAGAAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTCTGGCACTGATTTGA 180  
Qy 181 GAAAATAGAGAGCTTCGCACACTTTATGATTTATCTTTGCAATTTGATGGATTTGATGG 240  
Db 181 GAAAATAGAGAGCTTCGCACACTTTATGATTTATCTTTGCAATTTGATGGATTTGATGG 240  
Qy 241 AATCTCTTTACAGCAGATGAGAGGTTATATTAGTGGGCATAGTCTTGTTGGTGTATGAAT 300  
Db 241 AATCTCTTTACAGCAGATGAGAGGTTATATTAGTGGGCATAGTCTTGTTGGTGTATGAAT 300  
Qy 301 TGGGACTTGTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTCTGGCTGCTT 360  
Db 301 TGGGACTTGTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTCTGGCTGCTT 360  
Qy 361 TCATGCTGATTTCTGTTCAACAACCTCTCTCTTTGTTGGAACAGTATATGAGCGGACGC 420  
Db 361 TCATGCTGATTTCTGTTCAACAACCTCTCTCTTTGTTGGAACAGTATATGAGCGGACGC 420  
Qy 421 CAGCGGAGAAATGGTTGGATATCTAGTTTTTACCATATGTTTCCCTGAAAGCCACTGA 480  
Db 421 CAGCGGAGAAATGGTTGGATATCTAGTTTTTACCATATGTTTCCCTGAAAGCCACTGA 480  
Qy 481 CATCATGTTTTTTGGCCCAAGATTTCTGGCTCAAGCTCTACAGCTATGCTCTCTG 540  
Db 481 CATCATGTTTTTTGGCCCAAGATTTCTGGCTCAAGCTCTACAGCTATGCTCTCTG 540  
Qy 541 AGGATCTTCATTTAGCATCATCATGTTGAGACCAAGCTCTTTGTTATGGAAGCACTAT 600  
Db 541 AGGATCTTCATTTAGCATCATCATGTTGAGACCAAGCTCTTTGTTATGGAAGCACTAT 600  
Qy 601 CGAAGGCCAAGTATTTACAGATGAACGGTTTGGATCAAGTGAAGAGAGTTTACATTTGT 660  
Db 601 CGAAGGCCAAGTATTTACAGATGAACGGTTTGGATCAAGTGAAGAGAGTTTACATTTGT 660  
Qy 661 GCACTGAGATTAAGGCATACAGAGANTTCCAGGATGCGCAATTCACACATTTGGT 720  
Db 661 GCACTGAGATTAAGGCATACAGAGANTTCCAGGATGCGCAATTCACACATTTGGT 720  
Qy 721 TCACGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCTATGCGAGCCCAAA 780  
Db 721 TCACGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCTATGCGAGCCCAAA 780  
Qy 781 AACTTTGGCCCTCTCTTTGGAATTTGCCATAAATACAACTGATCTTACATTTATGCT 840  
Db 781 AACTTTGGCCCTCTCTTTGGAATTTGCCATAAATACAACTGATCTTACATTTATGCT 840  
Qy 841 TCGTCTCATGTCAAGATTTTCAAGTGCATGCTGATTTTCTATTTTTCGACCGGCG 900  
Db 841 TCGTCTCATGTCAAGATTTTCAAGTGCATGCTGATTTTCTATTTTTCGACCGGCG 900  
Qy 901 ATAACGTCTTTGCTTATTTAAGGATTCAGTAAATTTTCACTCTTCTAGTGGGAGGCT 960  
Db 901 ATAACGTCTTTGCTTATTTAAGGATTCAGTAAATTTTCACTCTTCTAGTGGGAGGCT 960  
Qy 961 TCACATAAGGATTTGTTCTTCTTCCATTCAGTGTGTTGTTATGTTGAGATCTTAAC 1020  
Db 961 TCACATAAGGATTTGTTCTTCTTCCATTCAGTGTGTTGTTATGTTGAGATCTTAAC 1020  
Qy 1021 CGTATCAATCTTGTATCAAACTCTCTCTTTTGGAAAAAAGAAAAA 1079  
Db 1021 CGTATCAATCTTGTATCAAACTCTCTCTTTTGGAAAAAAGAAAAA 1079

; Sequence 241778, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827,129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 241778  
; LENGTH: 607  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-241778

Query Match 2.7%; Score 29; DB 13; Length 607;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1051 TTCTTTTGGAAAAAAGAAAAA 1079  
Db 384 TTCTTTTGGAAAAAAGAAAAA 412

RESULT 3  
US-10-027-632-241778  
; Sequence 241778, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827,129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 241778  
; LENGTH: 607  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-241778

Query Match 2.7%; Score 29; DB 17; Length 607;  
Best Local Similarity 100.0%; Pred. No. 0.0041;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCTTTTGAAGAAAAA 1079  
DB 384 TTCTTTTGAAGAAAAA 412

## RESULT 4

US-10-425-115-79873/c  
; Sequence 79873, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 79873  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(615)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_17286C.1  
US-10-425-115-79873

Query Match 2.6%; Score 28; DB 20; Length 615;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TCCTTTTGAAGAAAAA 1079  
DB 68 TCCTTTTGAAGAAAAA 41

## RESULT 5

US-10-242-535A-17538  
; Sequence 17538, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17538  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-17538

Query Match 2.5%; Score 27; DB 17; Length 207;

Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
DB 154 CCTTTTGAAGAAAAA 180

## RESULT 6

US-10-085-783A-17538  
; Sequence 17538, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17538  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-17538

Query Match 2.5%; Score 27; DB 18; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
DB 154 CCTTTTGAAGAAAAA 180

## RESULT 7

US-09-983-965-4280/c  
; Sequence 4280, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 37-21(10297)C  
; CURRENT APPLICATION NUMBER: US/09/983,965  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 4280  
; LENGTH: 321  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 62-LIB3058-058-Q1-K1-H10  
US-09-983-965-4280

Query Match 2.5%; Score 27; DB 9; Length 321;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 47 CCTTTTGAAGAAAAA 21

## RESULT 8

US-09-873-367C-946/c  
; Sequence 946, Application US/09873367C  
; Publication No. US20030165839A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; APPLICANT: Soppet, Daniel  
; APPLICANT: Endress, Gregory  
; APPLICANT: Augustus, Meena  
; APPLICANT: Ebner, Reinhard  
; APPLICANT: Carter, Kenneth  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; TITLE OF INVENTION: Signature Gene Sets  
; FILE REFERENCE: 689290-64  
; CURRENT APPLICATION NUMBER: US/09/873,367C  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: U.S. 60/245,084  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 946  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-367C-946

Query Match 2.5%; Score 27; DB 10; Length 446;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 29 CCTTTTGAAGAAAAA 3

## RESULT 9

US-10-843-641A-946/c  
; Sequence 946, Application US/10843641A  
; Publication No. US2005006454A1  
; GENERAL INFORMATION:  
; APPLICANT: Avalon Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; TITLE OF INVENTION: Signature Gene Sets  
; FILE REFERENCE: 689290-189  
; CURRENT APPLICATION NUMBER: US/10/843,641A  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/873,367  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,832  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/964,824  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/09/967,768  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/09/968,007  
; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/09/969,347  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,708  
; PRIOR FILING DATE: 2001-10-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 8447  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 946  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-843-641A-946

Query Match 2.5%; Score 27; DB 21; Length 446;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 29 CCTTTTGAAGAAAAA 3

## RESULT 10

US-10-425-115-102591/c  
; Sequence 102591, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 102591  
; LENGTH: 511  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_25072C.1  
US-10-425-115-102591

Query Match 2.5%; Score 27; DB 20; Length 511;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 30 CCTTTTGAAGAAAAA 4

## RESULT 11

US-09-925-300-754  
; Sequence 754, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 754  
; LENGTH: 1795

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-754

Query Match          2.5%; Score 27; DB 9; Length 1795;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAAAGAAAAA 1079
    |||||||
Db 1757 CCTTTTGAAGAAAAAAGAAAAA 1783

RESULT 12
US-10-264-237-1253
; Sequence 1253, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn ver. 3.1
; SEQ ID NO 1253
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1253

Query Match          2.5%; Score 27; DB 17; Length 1795;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAAAGAAAAA 1079
    |||||||
Db 1757 CCTTTTGAAGAAAAAAGAAAAA 1783

RESULT 13
US-10-242-535A-54828
; Sequence 54828, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54828
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-54828

Query Match          2.4%; Score 26; DB 18; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||||
Db 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 15
US-10-437-963-72087/c
; Sequence 72087, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-54828

Query Match          2.4%; Score 26; DB 17; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||||
Db 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 14
US-10-085-783A-54828
; Sequence 54828, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54828
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-54828

Query Match          2.4%; Score 26; DB 18; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||||
Db 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 15
US-10-437-963-72087/c
; Sequence 72087, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72087
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72498C.1
US-10-437-963-72087

Query Match      2.4%; Score 26; DB 19; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTTGAAAAAAAAAAAAAAAAA 1079
      |||||||
Db 79 CTTTTGAAAAAAAAAAAAAAAAA 54

Search completed: September 12, 2005, 20:49:57
Job time : 737.243 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:35:08 ; Search time 204.853 Seconds  
(without alignments)  
8618.590 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaaagaaaagaaaa.....gaaaaaagaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	2.4	814	4	US-09-780-717-34
2	25	2.3	235	4	US-09-621-976-9455
3	25	2.3	492	4	US-09-270-767-6495
4	25	2.3	492	4	US-09-270-767-21777
5	25	2.3	1096	3	US-09-136-073-1
6	25	2.3	1096	3	US-09-457-024A-1
7	25	2.3	1811	2	US-08-808-931-9
8	25	2.3	1811	3	US-08-808-323-9
9	25	2.3	1811	3	US-09-050-603A-9
10	25	2.3	1811	3	US-09-102-420B-9
11	25	2.3	1811	3	US-09-497-698-9
12	25	2.3	1811	4	US-09-730-525-9
13	25	2.3	2080	4	US-09-311-021-179
14	25	2.3	3138	3	US-09-234-332-5
15	25	2.3	29930	4	US-09-949-016-15326
16	25	2.3	636591	4	US-09-949-016-11808
17	25	2.3	636591	4	US-09-949-016-11388
18	24	2.2	282	4	US-09-621-976-18648
19	24	2.2	299	4	US-09-621-976-7775
20	24	2.2	394	4	US-09-621-976-15204
21	24	2.2	517	4	US-10-101-464A-970
22	24	2.2	601	4	US-09-949-016-113460
23	24	2.2	1039	4	US-09-464-535-23
24	24	2.2	1140	2	US-08-598-805-5
25	24	2.2	1193	3	US-09-347-798-1
26	24	2.2	1699	3	US-09-152-060-19
27	24	2.2	1745	1	US-08-464-523B-1

ALIGNMENTS

RESULT 1

US-09-780-717-34  
; Sequence 34, Application US/09780717  
; Patent No. 6713666  
; GENERAL INFORMATION:  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Bates, Nic  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: No. 6713666el Invertase Inhibitors and Methods  
; TITLE OF INVENTION: of Use  
; FILE REFERENCE: 035718/208677  
; CURRENT APPLICATION NUMBER: US/09/780,717  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: US 60/181,509  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 814  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (99)...(638)  
US-09-780-717-34

Query Match 2.4%; Score 26; DB 4; Length 814;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1054 CTTTGTGAAAAA 1079  
|||||  
Db 770 CTTTGTGAAAAA 795

RESULT 2

US-09-621-976-9455  
; Sequence 9455, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9455

```
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9455

Query Match          2.3%; Score 25; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079
    |||
DB 176 TTTTGAAGAAAAAAGAAAAA 200

RESULT 3
US-09-270-767-6495
; Sequence 6495, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6495
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-6495

Query Match          2.3%; Score 25; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079
    |||
DB 385 TTTTGAAGAAAAAAGAAAAA 409

RESULT 4
US-09-270-767-21777
; Sequence 21777, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21777
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-21777

Query Match          2.3%; Score 25; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079
    |||
DB 385 TTTTGAAGAAAAAAGAAAAA 409

RESULT 5
US-09-136-073-1
; Sequence 1, Application US/09136073
; Patent No. 6043076
; GENERAL INFORMATION:
```

```
; APPLICANT: Ramakrishnan, Santha
; TITLE OF INVENTION: GENE ENCODING 2,3-DIHYDROXYBENZOIC ACID DECARBOXYLASE
; FILE REFERENCE: UNVN.53687
; CURRENT APPLICATION NUMBER: US/09/136,073
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,621
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Corel WordPerfect 6.1 for Windows
; SEQ ID NO 1
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: gene
; LOCATION: 1-1096
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1-1029
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 1073
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 1056-1061
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 1027-1073
; FEATURE:
; NAME/KEY: source/Aspergillus niger
; LOCATION: 1-1096
; PUBLICATION INFORMATION:
; AUTHORS: Santha, Ramakrishnan
; AUTHORS: Dickman, Martin B.
; AUTHORS: O'Leary, Marion H.
; TITLE: 2,3-Dihydroxybenzoic Acid Decarboxylase From Aspergillus niger:
; TITLE: Mechanism, Cloning And Overexpression.
; JOURNAL: Faseb Journal
; VOLUME: 11
; ISSUE: 9
; PAGES: A1017
; DATE: 1997-07-31
US-09-136-073-1

Query Match          2.3%; Score 25; DB 3; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079
    |||
DB 1071 TTTTGAAGAAAAAAGAAAAA 1095

RESULT 6
US-09-457-024A-1
; Sequence 1, Application US/09457024A
; Patent No. 6440704
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Santha
; TITLE OF INVENTION: GENE ENCODING 2,3-DIHYDROXYBENZOIC ACID DECARBOXYLASE
; FILE REFERENCE: UNVN.53687
; CURRENT APPLICATION NUMBER: US/09/457,024A
; CURRENT FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: 09/136,073
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Corel WordPerfect 6.1 for Windows
; SEQ ID NO 1
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: gene
```



```
; LOCATION: 1-1096
; NAME/KEY: CDS
; LOCATION: 1-1029
; NAME/KEY: polyA_site
; LOCATION: 1073
; NAME/KEY: polyA_signal
; LOCATION: 1056-1061
; NAME/KEY: 3' UTR
; LOCATION: 1027-1073
; NAME/KEY: source/Aspergillus niger
; LOCATION: 1-1096
; PUBLICATION INFORMATION:
; AUTHORS: Santha, Ramakrishnan
; AUTHORS: Dickman, Martin B.
; AUTHORS: O'Leary, Marion H.
; TITLE: 2,3-Dihydroxybenzoic Acid Decarboxylase From Aspergillus niger:
; TITLE: Mechanism, Cloning And Overexpression.
; JOURNAL: Faseb Journal
; VOLUME: 11
; ISSUE: 9
; PAGES: A1017
; DATE: 1997-07-31
; US-09-457-024A-1

Query Match      2.3%; Score 25; DB 3; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 1071 TTTTGGAAAAA 1095

RESULT 7
US-08-808-931-9
; Sequence 9, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1846
```

```
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pMDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
; US-08-808-931-9

Query Match      2.3%; Score 25; DB 2; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 1787 TTTTGGAAAAA 1811

RESULT 8
US-08-808-323-9
; Sequence 9, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,323
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1846
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
US-08-808-323-9

Query Match 2.3%; Score 25; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1055 TTTTGGAAAAAAGAAAAA 1079
Db 1787 TTTTGGAAAAAAGAAAAA 1811

RESULT 9
US-09-050-603A-9
; Sequence 9, Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphylinogen Oxidase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6023012artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,603A
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
;
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
US-09-050-603A-9

Query Match 2.3%; Score 25; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1055 TTTTGGAAAAAAGAAAAA 1079
Db 1787 TTTTGGAAAAAAGAAAAA 1811

RESULT 10
US-09-102-420B-9
; Sequence 9, Application US/09102420B
; Patent No. 6084155
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; TITLE OF INVENTION: OXIDASE ("PROTOX")
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6084155artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,420B
; FILING DATE: 22-JUN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/059,164
; FILING DATE: 13-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,603
; FILING DATE: 30-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/126,430
; FILING DATE: 11-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1811 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ORGANISM: Triticum aestivum (wheat)  
IMMEDIATE SOURCE:  
CLONE: PMDC-13 (NRRL B-21545)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1589  
OTHER INFORMATION: /product= "wheat protox-1"  
US-09-102-420B-9

Query Match 2.3%; Score 25; DB 3; Length 1811;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079  
|||||  
DB 1787 TTTTGAAGAAAAAAGAAAAA 1811  
|||||

RESULT 11  
US-09-497-698-9  
Sequence 9, Application US/09497698  
Patent No. 6308458  
GENERAL INFORMATION:  
APPLICANT: Volrath, Sandra  
Johnson, Marie  
Ward, Eric  
Heifetz, Peter  
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN  
OXIDASE ("PROTOX")  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6308458artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/497,698  
FILING DATE: 03-Feb-2000  
CLASSIFICATION: <Unknown>

30-MAR-1998  
11-MAR-1998  
28-FEB-1997  
28-FEB-1996  
28-FEB-1996  
21-JUN-1996  
06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/102,420  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 09/050,603  
FILING DATE: 30-MAR-1998  
APPLICATION NUMBER: US 60/126,430  
FILING DATE: 11-MAR-1998  
APPLICATION NUMBER: US 08/808,931  
FILING DATE: 28-FEB-1997  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1811 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Triticum aestivum (wheat)  
IMMEDIATE SOURCE:  
CLONE: PMDC-13 (NRRL B-21545)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1589  
OTHER INFORMATION: /product= "wheat protox-1"  
US-09-497-698-9

Query Match 2.3%; Score 25; DB 3; Length 1811;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079  
|||||  
DB 1787 TTTTGAAGAAAAAAGAAAAA 1811  
|||||

RESULT 12  
US-09-730-525-9  
Sequence 9, Application US/09730525  
Patent No. 6808904  
GENERAL INFORMATION:  
APPLICANT: Ward, Eric  
Volrath, Sandra  
Johnson, Marie  
Potter, Sharon  
TITLE OF INVENTION: Herbicide Tolerant Protox Genes  
Produced by DNA Shuffling  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6808904artis Corporation

```
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/730,525
FILING DATE: 05-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/102,419
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
APPLICATION NUMBER: US 09/038,878
FILING DATE: 11-MAR-1998
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1811 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Triticum aestivum (wheat)
IMMEDIATE SOURCE:
CLONE: pWDC-13 (NRRL B-21545)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1589
OTHER INFORMATION: /product= "wheat protox-1"
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-730-525-9
Query Match 2.3%; Score 25; DB 4; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1055 TTTTGAAGAAAAA 1079
Db 1787 TTTTGAAGAAAAA 1811
RESULT 13
US-09-311-021-179
; Sequence 179, Application US/09311021
; Patent No. 6706869
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
```

```
; APPLICANT: Fecthel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6300-11A
; CURRENT APPLICATION NUMBER: US/09/311,021
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 179
; LENGTH: 2080
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-311-021-179
Query Match 2.3%; Score 25; DB 4; Length 2080;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1055 TTTTGAAGAAAAA 1079
Db 2045 TTTTGAAGAAAAA 2069
RESULT 14
US-09-234-332-5/c
; Sequence 5, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; TITLE OF INVENTION: NEURONS; TRANSDIFFERENTIATION OF EPIDERMAL CELLS
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3138
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: gene
; LOCATION: {}...(0)
; OTHER INFORMATION: Zic 1 Protein gene; Genbank Acession D76435
US-09-234-332-5
Query Match 2.3%; Score 25; DB 3; Length 3138;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1054 CTTTGTGAAAAA 1078
Db 219 CTTTGTGAAAAA 195
RESULT 15
US-09-949-016-15326/c
; Sequence 15326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15326
; LENGTH: 29930
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(29930)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15326

```

```
Query Match      2.3%; Score 25; DB 4; Length 29930;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1055 TTTTGTGAAAAAAAAAAAAAAAAA 1079  
|||||  
Db 25446 TTTTGTGAAAAAAAAAAAAAAAAA 25422

Search completed: September 12, 2005, 17:29:48  
Job time : 207.853 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 08:41:37 ; Search time 3440.77 Seconds  
(without alignments)  
11936.677 Million cell updates/sec

Title: US-10-780-002-1

Perfect score: 1079

Sequence: 1 acgcgggaagaagaagaaaa.....gaaaaaagaaaaa 1079

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hic:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gse1:.\*  
9: gb\_gse2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	49	4.5	494	2	BE924354
2	41	3.8	384	2	BF097870
3	41	3.8	646	2	AW621893
4	34	3.2	420	1	AI772824
5	34	3.2	476	1	AI772823
6	34	3.2	535	2	BE433359
7	33	3.1	922	7	CK269872
8	31	2.9	525	1	AI775631
9	31	2.9	546	2	AW224223
10	31	2.9	590	4	BI928709
11	31	2.9	781	4	BG126704
12	30	2.8	124	2	BB583163
13	30	2.8	388	1	AI591310
14	30	2.8	400	1	AI718513
15	30	2.8	417	5	BQ383556
16	30	2.8	419	5	BQ159741
17	30	2.8	472	5	BY240192
18	30	2.8	535	4	BG804653
19	30	2.8	600	4	BG806457
20	30	2.8	600	5	BU924647
21	30	2.8	657	1	AL897127
22	30	2.8	684	6	CA316959
23	30	2.8	698	7	CN460297
24	30	2.8	855	5	BX683276

C	25	30	2.8	938	5	BU520408	BU520408	AGENCOURT
C	26	29	2.7	207	6	C93102	C93102	DICC
C	27	29	2.7	364	1	AI431179	8a22603.x	
C	28	29	2.7	399	7	CV496924	61077.1	M
C	29	2.7	548	6	CA814690	CA814690	CA48LN101	
C	30	29	2.7	614	7	CV258573	WS0249.B2	
C	31	29	2.7	966	7	CN644624	ILLUMIGEN	
C	32	28	2.6	218	2	BB589554	BB589554	
C	33	28	2.6	227	7	CO511445	T9E8Tzyq6	
C	34	28	2.6	294	2	BF651915	275331.MA	
C	35	28	2.6	434	4	BM392172	UI-R-D01-	
C	36	28	2.6	448	4	BI495836	df121b05.	
C	37	28	2.6	468	1	AI681388	tx46c02.x	
C	38	28	2.6	511	7	CK433247	UI-D-GC1-	
C	39	28	2.6	535	1	AI765652	w182g04.x	
C	40	28	2.6	550	7	CK433514	UI-D-GC1-	
C	41	28	2.6	561	7	CV225559	WS0161.B2	
C	42	28	2.6	581	7	CK433735	UI-D-GC1-	
C	43	28	2.6	622	7	CR585931	CR585931	
C	44	28	2.6	627	6	CA423834	UI-H-FE1-	
C	45	28	2.6	694	4	BI294974	UI-R-DK0-	

## ALIGNMENTS

RESULT 1  
BE924354  
LOCUS  
DEFINITION  
EST428123 potato leaves and petioles Solanum tuberosum cDNA clone  
CSTB29K16 5' sequence, mRNA sequence.  
ACCESSION  
BE924354  
VERSION  
BE924354.1 GI:10450430  
KEYWORDS  
EST.  
SOURCE  
Solanum tuberosum (potato)  
ORGANISM  
Solanum tuberosum  
REFERENCE  
1 (bases 1 to 494)  
AUTHORS  
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,  
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,  
Rønning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.  
TITLE  
Generation of ESTs from potato leaves and petioles  
JOURNAL  
Unpublished (2000)  
COMMENT  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/.

FEATURES  
source  
1..494  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cSTB29K16"  
/tissue\_type="leaflets and petioles"  
/dev\_stage="8 weeks old plants"  
/lab\_host="SOLR"  
/clone\_lib="potato leaves and petioles"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; Tissue was supplied by Dr. Fry (Cornell University).  
Leaflets and petioles were isolated from 8 week old  
greenhouse grown plants. The plants were watered and  
fertilized freely. The tissue was immediately frozen in  
liquid nitrogen."  
ORIGIN  
Query Match 4.5%; Score 49; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 8.9e-13;

```

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 TAGTGGGGCATAGTCTTGTTGGTATCAATTGGGACTTGCTATGGAAAA 319
|||||
Db 258 TAGTGGGGCATAGTCTTGTTGGTATCAATTGGGACTTGCTATGGAAAA 306
|||||

RESULT 2
BF097870 384 bp mRNA linear EST 18-MAY-2001
DEFINITION EST415943 tomato nutrient deficient roots Lycopersicon esculentum
CDNA clone CLW24G19 5' sequence, mRNA sequence.
BF097870
EST.
SOURCE BF097870.1 GI:10903580
ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 384)
REFERENCE van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
AUTHORS Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
Generation of ESTs from tomato nutrient-deficient roots
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..384
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="cLEW24G19"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/clone_lib="tomato nutrient deficient roots"
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="cLEW24G19"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/clone_lib="tomato nutrient deficient roots"

NOTE=Vector: pBluescriptSMCudapt; Site 1: 5' EcoRI;
Site 2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM Al, Zn,
P, K, Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."

ORIGIN
Query Match 3.8%; Score 41; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGTCTTGTTGGTATGAATTGGGACTTGCTATGGAAAA 319
|||||
Db 249 CATAGTCTTGTTGGTATGAATTGGGACTTGCTATGGAAAA 289
|||||

RESULT 3
AW621893 646 bp mRNA linear EST 18-MAY-2001
LOCUS EST312691 tomato root during/after fruit set, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cLEX13L15 5', mRNA sequence.
AW621893
EST.
SOURCE AW621893.1 GI:7333540
ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

```

```

REFERENCE 1 (bases 1 to 646)
AUTHORS van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Kochian,L.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato root, during and after fruit set
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..646
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX13L15"
/tissue_type="root"
/dev_stage="plants during and after fruit-set"
/clone_lib="tomato root during/after fruit set, Cornell
University"
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX13L15"
/tissue_type="root"
/dev_stage="plants during and after fruit-set"
/clone_lib="tomato root during/after fruit set, Cornell
University"
NOTE=Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: fanksley; tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."

ORIGIN
Query Match 3.8%; Score 41; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 8.5e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGTCTTGTTGGTATGAATTGGGACTTGCTATGGAAAA 319
|||||
Db 249 CATAGTCTTGTTGGTATGAATTGGGACTTGCTATGGAAAA 289
|||||

RESULT 4
AW621893 420 bp mRNA linear EST 18-MAY-2001
LOCUS EST253924 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION clone CLER4A4, mRNA sequence.
AW621893
EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 420)
REFERENCE D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
AUTHORS Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Other ESTs: TC2068
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..420
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER4A4"
/tissue_type="leaf"

```



```

/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLRR - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Query Match      3.2%; Score 34; DB 1; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361
|||||
DB 302 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 335

RESULT 5
AI772823      476 bp mRNA linear EST 18-MAY-2001
LOCUS EST253923 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION clone cLER4A2, mRNA sequence.
ACCESSION AI772823
VERSION AI772823
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 476)
AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Niemman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Other ESTs: TC2068
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source
Location/Qualifiers
1..476
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER4A2"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLRR - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Query Match      3.2%; Score 34; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361
|||||
DB 301 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 334

RESULT 6
AI772823
LOCUS EST253923 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION clone cLER4A2, mRNA sequence.
ACCESSION AI772823
VERSION AI772823
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 922)
AUTHORS Buell,C.R., Hart,A., Ziesmann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Other ESTs: EST715951
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org

```

```

LOCUS BE433359      535 bp mRNA linear EST 18-MAY-2001
DEFINITION EST399888 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
ACCESSION BE433359
VERSION BE433359
KEYWORDS EST.
SOURCE BE433359.1 GI:9431202
ORGANISM Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 535)
AUTHORS Alcalá,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niemman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source
Location/Qualifiers
1..535
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG13E6"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit, TIGR"
/notes="Vector: pBluescriptSKmCudapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopen accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN
Query Match      3.2%; Score 34; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361
|||||
DB 298 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 331

RESULT 7
CK269872      922 bp mRNA linear EST 03-AUG-2004
LOCUS EST715950 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION clone POACP55 5' end, mRNA sequence.
ACCESSION CK269872
VERSION CK269872
KEYWORDS EST.
SOURCE CK269872.1 GI:39826850
ORGANISM Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 922)
AUTHORS Buell,C.R., Hart,A., Ziesmann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Other ESTs: EST715951
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org

```

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

#### FEATURES

source

##### Location/Qualifiers

```
1. .922
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POACP55"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, 1 d,
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."
```

#### ORIGIN

Query Match 3.1%; Score 33; DB 7; Length 922;  
Best Local Similarity 100.0%; Pred. No. 8.1e-05;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 287 TGGTGGTATGAATTTGGGACTTGGCTATGGAAAA 319  
|||||  
Db 279 TGGTGGTATGAATTTGGGACTTGGCTATGGAAAA 311

#### RESULT 8

AI775631

##### LOCUS

DEFINITION EST256731 tomato resistant, Cornell Lycopersicon esculentum cDNA  
clone cLER1615, mRNA sequence.

##### ACCESSION

AI775631

##### VERSION

AI775631.1 GI:5273672

##### KEYWORDS

EST.

##### SOURCE

Lycopersicon esculentum (tomato)

##### ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 525)

D' Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,

Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,

Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,

Giovannoni, J.J., and Martin, G.B.

Generation of ESTs from Pseudomonas resistant tomato

Contact: CUGI

Unpublished (1999)

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

#### FEATURES

source

##### Location/Qualifiers

```
1. .525
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER1615"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."
```

#### ORIGIN

Query Match 2.9%; Score 31; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.00082;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 CATAGTCTTGCTGATGAATTTGGACTTG 309  
|||||  
Db 274 CATAGTCTTGCTGATGAATTTGGACTTG 304

#### RESULT 9

AW224223

##### LOCUS

DEFINITION AW224223 546 bp mRNA linear EST 18-MAY-2001

clone cLEN16B20, mRNA sequence.

##### ACCESSION

AW224223

##### VERSION

AW224223.1 GI:5536003

##### KEYWORDS

EST.

##### SOURCE

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 546)

Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,

Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S.,

Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and

Giovannoni, J.

Generation of ESTs from tomato fruit tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

##### Location/Qualifiers

```
1. .546
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN16B20"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni; Fruit were tagged at the
breaker stage (first sign of lycopene accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
```

#### ORIGIN

Query Match 2.9%; Score 31; DB 2; Length 546;  
 Best Local Similarity 100.0%; Pred. No. 0.00082;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309  
 |||||  
 Db 358 CATAGCTTGGTGTATGAATTTGGGACTTG 388

RESULT 10  
 BI928709  
 LOCUS  
 DEFINITION EST548598 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA clone CT0B26K14 5' end, mRNA sequence.

ACCESSION  
 VERSION BI928709.1 GI:16240757  
 KEYWORDS EST.

SOURCE  
 ORGANISM Lycopersicon esculentum (tomato)

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 590)  
 van der Hoeven, R.S., Bezzerides, J.L., Karanycheva, S.A., Tsai, J., Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.  
 Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
 Unpublished (2001)

CONTACT: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics Institute  
 Seq primer: T3.

FEATURES  
 source  
 Location/Qualifiers  
 1..590  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CT0B26K14"  
 /tissue\_type="flower"  
 /dev\_stage="3-8mm buds"  
 /clone\_lib="tomato flower, 3 - 8 mm buds"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

## ORIGIN

Query Match 2.9%; Score 31; DB 4; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 0.00082;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309  
 |||||  
 Db 272 CATAGCTTGGTGTATGAATTTGGGACTTG 302

## RESULT 11

BI928704

LOCUS  
 DEFINITION EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone CT0F13K13 5' sequence, mRNA sequence.

ACCESSION  
 VERSION BI928704.1 GI:12626892  
 KEYWORDS EST.

SOURCE  
 ORGANISM Lycopersicon esculentum (tomato)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 781)  
 van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T., Hansen, C., Ronning, C. and Tanksley, S.  
 Generation of ESTs from tomato shoot/meristem tissue  
 Unpublished (2001)

CONTACT: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>.

## FEATURES

source

Location/Qualifiers  
 1..781  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CT0F13K13"  
 /tissue\_type="shoot/meristem"  
 /dev\_stage="developing shoots from 4-6wks old plants"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato shoot/meristem"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA496). Tissue was immediately frozen in liquid nitrogen."

## ORIGIN

Query Match 2.9%; Score 31; DB 4; Length 781;  
 Best Local Similarity 100.0%; Pred. No. 0.00081;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309  
 |||||  
 Db 282 CATAGCTTGGTGTATGAATTTGGGACTTG 312

## RESULT 12

BI928704

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BI928704  
 BI928704 RIKEN full-length enriched, adult male cecum Mus musculus cDNA clone 9130204I12 5', mRNA sequence.

BI928704  
 BI928704.1 GI:11479707  
 EST.

Mus musculus

Mus musculus

Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 124)

Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno, H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Aizawa, K. et al. 2000)

Unpublished (2000)

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: [genome-res@gsr.riken.jp](mailto:genome-res@gsr.riken.jp), URL: <http://genome.gsc.riken.jp/>

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S.,

Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
Okazaki,Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

#### FEATURES

Location/Qualifiers  
1. .124  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="9130204112"  
/sex="male"  
/tissue\_type="cecum"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, adult male cecum"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGAGTCGAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 185.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGAGATTCGAGTAAATTAATTAATTCCTCCCTCCCTCC  
3']. cDNA was cloned into the XhoI and BamHI sites.  
Vector: a modified pBluescript KS(+) after bulk excision  
from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:  
BamHI"

#### ORIGIN

Query Match 2.8%; Score 30; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCTTTTGGAAAAAATTTTAAAAA 1079  
|||||  
Db 58 TTTCCTTTTGGAAAAAATTTTAAAAA 29  
|||||

#### RESULT 13

AI591310/c  
LOCUS  
DEFINITION  
tw11f05.x1 NCI CGAP Brn52 Homo sapiens cDNA clone IMAGE:2259393 3',  
similar to contains\_MER22.b3 MER22 repetitive element ;, mRNA  
sequence.

ACCESSION  
AI591310  
VERSION  
AI591310.1 GI:4600358  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 388)

REFERENCE  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BIAGP), Tumor Gene Index

#### JOURNAL COMMENT

Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 1681 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 388  
POLYA=No.

#### FEATURES

Location/Qualifiers  
1. .388  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2259393"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Brn52"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; This library represents the normalized  
version of NCI CGAP Brn35. Cloned unidirectionally.  
Primer: Oligo dT. Average insert size 1.19 kb. Tumor  
types include: meningioma, oligodendroglioma, astrocytoma  
(grade II), medulloblastoma, astrocytoma (grade IV).  
Constructed by Life Technologies."

#### ORIGIN

Query Match 2.8%; Score 30; DB 1; Length 388;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCTTTTGGAAAAAATTTTAAAAA 1079  
|||||  
Db 89 TTTCCTTTTGGAAAAAATTTTAAAAA 60  
|||||

#### RESULT 14

AI1718513/c  
LOCUS  
DEFINITION  
as56h02.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone  
IMAGE:2332755 3', similar to SW:CA11\_RAT P02454 COLLAGEN ALPHA 1(I)  
CHAIN ;, mRNA sequence.

ACCESSION  
AI1718513  
VERSION  
AI1718513.1 GI:5035769  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 400)

REFERENCE  
AUTHORS  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-NCI human EST Project  
Unpublished (1997)

#### TITLE JOURNAL COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 388.

FEATURES  
source

Location/Qualifiers  
1. .400  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2332755"  
/sex="male"  
/dev stage="adult, age 25"  
/lab host="DH10B (phage resistant)"  
/clone lib="Barstead colon HPLR87"  
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTGAATCTCAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."

ORIGIN

Query Match 2.8%; Score 30; DB 1; Length 400;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCCTTTTGGAAAAA 1079  
|||||  
DB 88 TTTCCCTTTTGGAAAAA 59

RESULT 15

BQ383556/c  
LOCUS BQ383556 417 bp mRNA linear EST 22-MAY-2002  
DEFINITION NISC\_mn01h05.x1 NICHD\_XGC\_Ov1 Xenopus laevis cDNA clone  
IMAGE:5049056 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Xenopus; Xenopus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.  
National Institute of Child Health and Human Development, National  
Cancer Institute, Xenopus Gene Collection  
Unpublished (2002)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
cDNA Library Preparation:  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
info@image.llnl.gov  
Plate: LHAM1133 row: 0 column: 9  
Seq primer: -21M13 forward primer (AB1).

FEATURES

source

Location/Qualifiers  
1. .417  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:5049056"  
/sex="female"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NICHD\_XGC\_Ov1"  
/note="Organ: ovary; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.0 kb. Constructed by Life  
Technologies."

ORIGIN

Query Match 2.8%; Score 30; DB 5; Length 417;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCCTTTTGGAAAAA 1079  
|||||  
DB 107 TTTCCCTTTTGGAAAAA 78

Search completed: September 12, 2005, 17:23:24  
Job time : 3449.77 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 06:32:02 ; Search time 590.713 Seconds  
(without alignments)  
10813.032 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1079	100.0	1079	8 ADA26498	Ada26498 Tobacco s
2	37	3.4	121	12 ADG00027	Adg00027 Nicotiana
C 3	27	2.5	258	2 AAQ61367	Aaq61367 Human bra
C 4	27	2.5	321	8 ABX54351	Abx54351 Bovine ES
C 5	27	2.5	446	6 ABL62609	AbL62609 Colon ade
6	27	2.5	552	8 ABZ73132	Abz73132 Rice leaf
7	27	2.5	1795	3 AAF16319	Aaf16319 Human pro
8	27	2.5	1795	6 ABL90691	AbL90691 Human pol
9	26	2.4	342	3 AAC79879	Aac79879 Human sec
10	26	2.4	375	3 AAC79873	Aac79873 Human sec
11	26	2.4	375	4 AAF26589	Aaf26589 DNA encod
12	26	2.4	814	5 AAS11362	Aas11362 Soybean D
13	26	2.4	1099	10 ADA13386	Ada13386 Human int
14	26	2.4	1664	2 AAQ73389	Aaq73389 Rabies vi
15	26	2.4	1870	2 AAQ73371	Aaz06371 Human rec
16	26	2.4	2243	3 AAC59832	Aac59832 Human sec
C 17	26	2.4	3820	13 ADR85197	Adr85197 Aspergill
C 18	26	2.4	5820	13 ADR84610	Adr84610 Aspergill
C 19	26	2.4	20752	4 AAK75098	Aak75098 Human inm
20	26	2.4	98345	13 ABD32892	Abd32892 Human can

C 21	25	2.3	148	5 ABV60761	Abv60761 Human pro
C 22	25	2.3	158	5 ABV61010	Abv61010 Human pro
C 23	25	2.3	194	5 ABV60879	Abv60879 Human pro
C 24	25	2.3	214	8 ABX50662	Abx50662 Bovine ES
25	25	2.3	219	6 ABX29070	Abx29070 cDNA enco
26	25	2.3	219	6 ABS68210	AbS68210 cDNA enco
27	25	2.3	219	10 ADC25204	Adc25204 Human cDN
C 28	25	2.3	221	5 ABV61804	Abv61804 Human pro
29	25	2.3	244	12 ADO41106	Ado41106 Human cDN
30	25	2.3	245	2 AAV30925	Aav30925 Human sec
31	25	2.3	245	5 AAF98404	Aaf98404 3' portio
C 32	25	2.3	276	12 ADM91205	Adm91205 Human DNA
33	25	2.3	282	11 ADT96541	Adt96541 Colon can
C 34	25	2.3	290	5 ABV61419	Abv61419 Human pro
C 35	25	2.3	310	5 ADQ18503	Adq18503 Human sof
C 36	25	2.3	341	5 ADI71926	Adi71926 Human ova
C 37	25	2.3	341	5 ADL37077	Adl37077 Human ova
38	25	2.3	343	4 AAD05605	Aad05605 Human sec
C 39	25	2.3	372	6 ABQ58882	Abq58882 Human col
C 40	25	2.3	393	5 ABV61180	Abv61180 Human pro
C 41	25	2.3	403	4 AAI82336	Aai82336 Human pol
C 42	25	2.3	409	9 ACH45796	Ach45796 Human foe
C 43	25	2.3	419	8 ABX48773	Abx48773 Bovine ES
44	25	2.3	425	4 AAI92294	Aai92294 Human pol
C 45	25	2.3	435	9 ACH20634	Ach20634 Human adu

ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
FT CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
XX 27-FEB-2003.  
XX  
PD 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
XX Klessig DF, Kumar D;  
XX WPI: 2003-278578/27.  
XX P-PSDB; ADA26499.

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.  
XX  
XX Claim 1; Fig 4; 98pp; English.  
XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the SABP2 protein from tobacco plants.

XX Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1079; DB 8; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGGAGAAAGAAACAAAGGCAATAAAATTCAAATGAAGGAAGAAACACT 60  
DB 1 ACGCGGGAGAAAGAAACAAAGGCAATAAAATTCAAATGAAGGAAGAAACACT 60  
QY 61 TTGTTTTAGTACATGCTGTCATGCCATGCGAGGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
DB 61 TTGTTTTAGTACATGCTGTCATGCCATGCGAGGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
QY 121 TAGAAGCTCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTCTTGCCACTGATTTGA 180  
DB 121 TAGAAGCTCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTCTTGCCACTGATTTGA 180  
QY 181 GAAATATAGAGAGCTTCGCACACTTTATGATTAATCTTTGCGCATGATGGATGG 240  
DB 181 GAAATATAGAGAGCTTCGCACACTTTATGATTAATCTTTGCGCATGATGGATGG 240  
QY 241 AATCTCTTTACAGATGAGAGGTTATATTAGTGGGCATAGTCTTGCTGATGAATT 300  
DB 241 AATCTCTTTACAGATGAGAGGTTATATTAGTGGGCATAGTCTTGCTGATGAATT 300  
QY 301 TGGGACTTCTGATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
DB 301 TGGGACTTCTGATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
QY 361 TCATGCTGATTTCTGTTCAACTCTCTCTTGTGTTTGGAAACAGTATAATGAGCGGAC 420  
DB 361 TCATGCTGATTTCTGTTCAACTCTCTCTTGTGTTTGGAAACAGTATAATGAGCGGAC 420  
QY 421 CAGCCGAGAAATGGTTGGATACCTCAGTTTTTACCATATGTTCCCTGAAGGCCACTGA 480  
DB 421 CAGCCGAGAAATGGTTGGATACCTCAGTTTTTACCATATGTTCCCTGAAGGCCACTGA 480  
QY 481 CATCCATGTTTTTGGCCCAAAAGTTCTTGGCTCACAAAGCTTACCAAGCTATGCTCTCCTG 540  
DB 481 CATCCATGTTTTTGGCCCAAAAGTTCTTGGCTCACAAAGCTTACCAAGCTATGCTCTCCTG 540  
QY 541 AGGATCTTGCATAGCATCATCATGTTGGTGAGACCAAGCTCTTGTATTGGAAGCCTAT 600  
DB 541 AGGATCTTGCATAGCATCATCATGTTGGTGAGACCAAGCTCTTGTATTGGAAGCCTAT 600  
QY 601 CGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGCT 660  
DB 601 CGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGCT 660  
QY 661 GCATGAGGATAAGGCATACCAAGAAATTCACGCGATGGCAAAATTCACATTTGGTG 720  
DB 661 GCATGAGGATAAGGCATACCAAGAAATTCACGCGATGGCAAAATTCACATTTGGTG 720  
QY 721 TCATGAGCAATAGAGATTAAGTGCTGATCACATGCAATGCTATGCGAGCCCAAA 780  
DB 721 TCATGAGCAATAGAGATTAAGTGCTGATCACATGCAATGCTATGCGAGCCCAAA 780  
QY 781 AACTTTGCGCTCTCTCTTGGAAATGGCCCATAAATACAACTGATCTTACATTATGCT 840  
DB 781 AACTTTGCGCTCTCTCTTGGAAATGGCCCATAAATACAACTGATCTTACATTATGCT 840

DB 781 AACTTTGCGCTCTCTCTTGGAAATGGCCCATAAATACAACTGATCTTACATTATGCT 840  
QY 841 TCGTCTCATGTCAGATTTTCAGTGCATGCTGTAATTTTTTTTCTATTTTCGACCGGCG 900  
DB 841 TCGTCTCATGTCAGATTTTCAGTGCATGCTGTAATTTTTTTTCTATTTTCGACCGGCG 900  
QY 901 ATAACTGCTTTGGCTATTTTAAGGATTCAGATTAATTTCACTCTTCTAGTGTGAAGGCT 960  
DB 901 ATAACTGCTTTGGCTATTTTAAGGATTCAGATTAATTTCACTCTTCTAGTGTGAAGGCT 960  
QY 961 TCCCATGAAGGATGTTCTGTTTCTCCATTCAAGTGTGTTATGTTGAGATACTTAAAC 1020  
DB 961 TCCCATGAAGGATGTTCTGTTTCTCCATTCAAGTGTGTTATGTTGAGATACTTAAAC 1020  
QY 1021 CGTATCAATCTTGTATGAATGAAACTTCTTCTTTTCTTTTGAAAAAATAAAAAA 1079  
DB 1021 CGTATCAATCTTGTATGAATGAAACTTCTTCTTTTCTTTTGAAAAAATAAAAAA 1079

RESULT 2  
ADG00027

ID ADG00027 standard; cDNA; 121 BP.

XX AC ADG00027;

XX 26-FEB-2004 (first entry)

XX Nicotiana tabacum variant bright yellow nucleotide sequence SEQ ID:544.  
XX secondary metabolite modulator; biosynthesis; alkaloid; phenylpropanoid;  
KW tobacco; Nicotiana tabacum variant bright yellow; Nicotiana tabacum BY;  
KW plant; gene; ss.

XX Nicotiana tabacum.

XX WO2003097790-A2.

XX 27-NOV-2003.

XX 16-MAY-2003; 2003WO-EP050171.

XX 17-MAY-2002; 2002EP-00076973.

XX 04-JUL-2002; 2002EP-00077674.

XX (VLAA-) VLAAIS INTERUNIVERSITAIR INST BIOTECHNOG.

XX (VTTB-) VTT BIOTECHNOLOGY.

XX Inze DG, Goossens A, Oksman-Caldentey K, Haekkinen ST, Laakso IJ;

XX WPI; 2004-022853/02.

XX New isolated polypeptides and polynucleotides useful for modulating the

XX biosynthesis of secondary metabolites (e.g. alkaloids or

XX phenylpropanoids) in an organism or its cell.

XX Claim 3; SEQ ID NO 544; 140pp; English.

XX The present invention describes an isolated polypeptide that modulates  
XX the production of at least one secondary metabolite in an organism, or  
XX its derived cell. The polypeptide is selected from a polypeptide encoded  
XX by a polynucleotide comprising any of the 871 nucleotide sequences given  
XX in the specification. The polypeptides and polynucleotides are useful in  
XX modulating the biosynthesis of secondary metabolites (e.g. alkaloids or  
XX phenylpropanoids) in an organism or its derived cell. The present  
XX sequence is used in the exemplification of the present invention.

XX Sequence 121 BP; 44 A; 19 C; 31 G; 27 T; 0 U; 0 Other;

Query Match 3.4%; Score 37; DB 12; Length 121;  
Best Local Similarity 100.0%; Pred. No. 0.00077;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628 GGTTTGGATCAGTGAAGAGATTTACATTGTGTGCAC 664



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 22:54:15 ; Search time 4681.27 Seconds  
(without alignments)  
11168.596 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaaagaaaa.....gaaaaaaagaaaaaa 1079

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sbs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

ALIGNMENTS  
AY485932 1079 bp mRNA linear PLN 07-JAN-2004  
Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.  
ACCESSION AY485932  
VERSION AY485932.1 GI:40549302  
KEYWORDS  
SOURCE Nicotiana tabacum (common tobacco)  
ORGANISM Nicotiana tabacum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiids; Solanales; Solanaceae; Nicotiana.

REFERENCE  
1 (bases 1 to 1079)  
Kumar, D. and Klessig, D.F.  
High-affinity salicylic acid-binding protein 2 is required for  
plant innate immunity and has salicylic acid-stimulated lipase  
activity  
Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)  
PUBMED 14673096  
REFERENCE 2 (bases 1 to 1079)  
Kumar, D. and Klessig, D.F.  
Direct Submission  
Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA  
JOURNAL  
FEATURES  
Location/Qualifiers  
1..1079  
/organism="Nicotiana tabacum"  
/mol\_type="mRNA"  
/cultivar="Xanthi nc [NN]"  
/db\_xref="taxon:4097"  
42..824  
/note="lipase; SAMP2"  
/codon\_start=1  
/product="salicylic acid-binding protein 2"  
/protein\_id="AA87711.1"  
/translation="MKEGKHFLVHGACHGWSYKLPLEAAGHKVTALDLAASGT  
DLRKTLEELTYDYTLPLMELMESLSADEKVTLVGLSGMNLGAMKEYPKIYAAV  
FLAAMPFVSNNSSFLVQYNERTPAENWLDQFLPYGSPPEPLSPFLAHLK  
YOLCSFEDLALASSLVRPSSFMDLSKAKYFTDRFGSVKRVYIVCTSDKGIPEEFQ  
RWQIDNIGVTEAIEIKGADHMAWLCEPQKLCASLLEIAHKYN"

ORIGIN  
BT006227 Arabidops  
AX506583 Sequence  
AX651280 Sequence  
AY046015 Arabidops  
AY142490 Arabidops  
AY093714 Arabidops  
AY010709 Arabidops  
AY058115 Arabidops  
BT012414 Arabidops  
AY086981 Arabidops  
AY074858 Arabidops  
AY142031 Arabidops  
AY455313 Lycopersi  
BT012867 Lycopersi  
AK061058 Oryza sat  
U40402 Hevea bras  
A59586 Sequence 1  
Z29091 M.esculent  
CQ829357 Sequence  
AX654444 Sequence  
CQ829358 Sequence  
AC003040 Arabidops  
AC004482 Arabidops  
AL161590 Arabidops  
Z99707 Arabidops  
AJ223281 Manihot e

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length	Description
1	1079	100.0	1079	AY485932 Nicotiana
2	352.6	32.7	1077	AF269158 Citrus si
3	332.8	30.8	961	AF178576 Rauwolfia
4	312.4	29.0	891	AF751530 Catharan
5	295.2	27.4	935	BT002859 Arabidops
6	292.2	27.1	792	AX506625 Sequence
7	292.2	27.1	792	BT014881 Arabidops
8	273	25.3	870	AY086590 Arabidops
9	272.6	25.3	922	AF361627 Arabidops
10	271	25.1	792	AX506856 Sequence
11	271	25.1	792	AY056068 Arabidops
12	269	24.9	792	CQ806104 Sequence
13	269	24.9	792	AX505522 Sequence
14	269	24.9	792	BT015031 Arabidops
15	269	24.9	878	BT014901 Arabidops
16	261	24.2	902	AK175772 Arabidops
17	256	23.7	792	BT010429 Arabidops
18	251	23.3	924	BT004125 Arabidops
19	244.8	22.7	928	AK117107 Arabidops



Db 347 AATTCACACAAAAATCTCCGTGGCTGTTTTCGTAACTGCATTCATGCCTGCACACACAC 406

QY 379 ACAACTCTCTCTTTGTTTGGACAGTATATAATCAGCGGAGCCCA-----GCCGGAAT 432

Db 407 ACCGGCAATCTTTTGGACAGTATTTCTGAGAAGATGGGAAAAGAGGACGACGCT 466

QY 433 GGTGGATACCTAGTTTACCATATGTTTCCCTGAAAGGCCACTGACATCATGTTT 492

Db 467 GGTGGACACTCAATTTTCAATGTGACGCGTCAAAATCCATCTCACATTTCCATGCTTT 526

QY 493 TTGGCCCAAGTTCTTGGCTCAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCAT 552

Db 527 TCGGCGCGAGTTCTTGAATATCAAGATCTATCAGCTTTGTCTCTCTGAGGATCTGGAGC 586

QY 553 TAGCATCATCATTTGGTGGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAGT 612

Db 587 TGCCCAAGATGTTTGGTGAGCCAGGATCAATGTTTATAGCAACTTATCGAAGAAAGTA 646

QY 613 ATTTCAAGATGAACGGTTTGGATCAGTGAAGAGATTTTACATTTGTGCTGACTGAGGATA 672

Db 647 AGTTCAAGATGAAGATACGGATCTGTTAAGCGAGTTTATCTTGTATGGAAGAGGATA 706

QY 673 AAGCATACCAAGAAATTCACGAGTGGCAATTCACAACATTTGTGTCTACATGAAGCAA 732

Db 707 TTGGTCTCCCTAAGCAATTTACGACTGGATGATCCAAAACCTATCCCGTTAATGAAGTGA 766

QY 733 TAGAGATTAAGGTGCTGATCAGATGCAATGCTATGCGAGCCGCCAAAACCTTTGCGCCT 792

Db 767 TGAGATCAAGGCGGTGATCAGTGGCAATGCTTTCCGACCCACAGAAACCTTTGCGATT 826

QY 793 CTCTCTTGGAAATGGCCATAATA 817

Db 827 GTCGTCTCAGATTTCTCTTAAGTA 851

RESULT 3

AF178576

LOCUS

DEFINITION

AF178576

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AF178576 961 bp mRNA linear PLN 29-MAR-2000

Rauvolfia serpentina polynuridine aldehyde esterase mRNA, complete cds.

AF178576 GI:6651392

Rauvolfia serpentina (serpentwood)

Rauvolfia serpentina

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae; Rauvolfia.

Dogru, E., Warzecha, H., Seibel, F., Haebel, S., Lottspeich, F. and Stockigt, J.

The gene encoding polynuridine aldehyde esterase of monoterpene indole alkaloid biosynthesis in plants is an ortholog of the alpha/betahydrolase super family

Eur. J. Biochem. 267 (5), 1397-1406 (2000)

20156243

10691977

2 (bases 1 to 961)

Warzecha, H., Dogru, E. and Stockigt, J.

Direct Submission

Submitted (18-AUG-1999) Institut fur Pharmazie, Johannes Gutenberg Universitaet, Staudinger Weg 5, Mainz 55099, Germany

Location/Qualifiers

1. .961

/organism="Rauvolfia serpentina"

/mol\_type="mRNA"

/db\_xref="taxon:4060"

/notes="cell suspension culture"

15. .809

/function="catalyzes 16-epi-velllosimine formation from polynuridine aldehyde"

/function="involved in ajmaline biosynthesis"

/codon\_start=1

/product="polynuridine aldehyde esterase"

/protein\_id="AAP22288.1"

/db\_xref="GI:6651393"

/translation="MHSAANAKQKHFLVHGGCLGANIYWKLPKLLSAGAHKTVADVLSAAGINRRRLDEIHTFDYSEPLMEVNASIPPEKVVLLHSGFSGMISGLAMETYPEKISAVFMSAMPDPNHSILTYPEKYNEKCPADMLDQSFTSTYNGPFGMISGLPQFMALXMFONCSVEDLELAKMLTRPGSLFQDLAKAKKFSERYGVSKRAYIFCNEDKSPFVFQKWFVESVGADKVKIEKADHMGMLSQPREVCKLLDISDS"

ORIGIN

Query Match 30.8%; Score 332.8; DB 8; Length 961;

Best Local Similarity 64.3%; Pred. No. 1.1e-72;

Matches 499; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 45 AAGGAGGAAACACATTTGTTTGTACATGTCGATGTCATGAGGTTTGGATGGTAC 104

Db 36 AAGCAACAAACATTTTGTCTGTACACGGCGGATGTCTCGAGCTTGGATCTGGTAC 95

QY 105 AAGCTAAAGCCACTGCTAGAACCTGCAGGCCATAAGGTTTACAGCCCTTGATTTAGCAGCT 164

Db 96 AAGCTCAAGCCGCTGCTCGAGTCAGCCGACATAGGTCAACCGCTTGACCTGTCGGCC 155

QY 165 TCTGGCACTGATTTGAGAAAAATAGAGGCTTCGCACACTTTATGATATATATTGCGCA 224

Db 156 GCCGCGCATCAACCCAAAGAGGCTCGATGAGATTCAACATTTCCGGACTACTCTGGAGCCC 215

QY 225 TTGATGGAGTTGATGGAATCTCTTTTCAGCAGATGAGAGGTTTATATTAGTGGGCATAGT 284

Db 216 TTGATGGAGTTGATGGAATCTCTCTCTGATGAGAGGTTTGTCTCTTCTGGCCATAGC 275

QY 285 CTGTTGGTGTATGAATTTGGGACTTCTGATGGAAGATATCCACAAAAGATCTATGCTGCT 344

Db 276 TTGTTGGCATGAGTTTGGTCTTCCCATGGAAACCTTCCAGAGAGATATACAGTTGCT 335

QY 345 GTTTTCTTGGCTGCTTTTCATGCTGATTTCTGTTTCAAACTCCTCTCTCTTTTGGAAACAG 404

Db 336 GTTTTATGTCTGCAATGATGCTCTTAACCACTCACTAACTTATCCGTTTGAGAAG 395

QY 405 TATATAGCGGACGCCAGCGAGAAATGGTTGGATCTCAGTTTTTACCATATGTTTCC 464

Db 396 TACAATGAGAAGTGTCCGCGAGATATGATGTTGGACTCACAGTTTTCACCTACGGAAC 455

QY 465 CCGTGAAGAGCCACTGCACATCCATGTTTGTGGCCCAAGGTTCTTGCTCACAAGCTCTAC 524

Db 456 CAGAGAACCCAGGAATGTCAATGATTTTGGAACCTCAGTTTATGGCCCTCAAAATGTTTC 515

QY 525 CAGCTATGCTCTCTGAGGATCTTGCATTTAGCATCATCATTTGGTGAGACCAAGCTCTTTG 584

Db 516 CAGAATTTGCTCAGTCAGGACCTTGAATTAGCCAAAATGTTGACTCGACCCAGGTTCTGTTA 575

QY 585 TTTATGGAAGACCTATCGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAG 644

Db 576 TTTTTCGAAGATTTGGCCAAAGCCAAAAGTTTCTCAACGAGAGGATTCGTTTCGGTGAAG 635

QY 645 AGAGTTTACATTTGCTGCTGACATGAAGGATTAAGGATACACAGAGAGATTCACGAGTGGCAA 704

Db 636 CAGCTTATATCTTTTGAATGAAGATAATCAATTTCCAGTTGAGTTTCAGAAATGTTT 695

QY 705 ATTGAACAATTTGGTGTCTCACTGAAGCAATAGAGATTTAAAGGTGCTGATCAGATGGCAATG 764

Db 696 GTTGAAGTGTGGAGCTGATAAAGTAAAGAAATCAAGAGAGCAGATCATATGGGAATG 755

QY 765 CTATCGGAGCCCCAAAAACCTTTGGCCCTCTCTCTTTGGAAATTTGCCCAATAACAA 820

Db 756 CTTTCGACGCAAGGGAAGTTTGCAAGTGCCTGCTTGTATATATATATATATATATATAT 811

RESULT 4

AY751530

LOCUS

DEFINITION

ACCESSION

AY751530

AY751530 891 bp mRNA linear PLN 11-OCT-2004

Catharanthus roseus protein S mRNA, complete cds.

```
VERSION      AY751530.1  GI:53830669
KEYWORDS
SOURCE      Catharanthus roseus (Madagascar periwinkle)
ORGANISM    Catharanthus roseus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiales; Gentianales; Apocynaceae; Rauvolfioideae;
            Vinceae; Catharanthus.
REFERENCE    1 (bases 1 to 891)
AUTHORS     Lemenager,D., Rideau,M. and Clastre,M.
TITLE       Cloning of the Catharanthus roseus protein S (CrPS) associated with
            monoterpenoid indole alkaloid production
JOURNAL     Unpublished
REFERENCE    2 (bases 1 to 891)
AUTHORS     Lemenager,D., Rideau,M. and Clastre,M.
TITLE       Direct Submission
JOURNAL     Submitted (15-SEP-2004) Laboratoire de Biologie Molculaire et
            Biochimie Vegetale, Faculte de Pharmacie, 31 Avenue Monge, Tours
            37200, France
FEATURES
            source
            1..891
            /organism="Catharanthus roseus"
            /mol_type="mRNA"
            /db_xref="taxon:4058"
            48..824
            /note="hydrolase; CrPS; alpha/beta fold family protein"
            /codon_start=1
            /product="protein S"
            /protein_id="AAU95203.1"
            /db_xref="GI:53830670"
            /translation="MEVVKHFVTVHGVHGAWVYKLPKTEAAGHRCITAVNLAAGSI
            NEKKUEVRSSIDYAPLLEVDVSPENEKVLVGHVGGGWTAAVGMKFPNKISLAV
            FLNAMPDTEPNSVLYEATTPPEAKDKQFSAYGDPPTISLVCGPFISLTYH
            LSPIDHAGKLIVRPGSLIFIEDLKAKEFTBEGFGSPRVVYVIAEDKTIPTPEFRW
            MEENPVKEVKEIKGADHMPSPKDELSQLLDIAKKA"
CDS
            1..891
            /note="hydrolase; CrPS; alpha/beta fold family protein"
            /codon_start=1
            /product="protein S"
            /protein_id="AAU95203.1"
            /db_xref="GI:53830670"
            /translation="MEVVKHFVTVHGVHGAWVYKLPKTEAAGHRCITAVNLAAGSI
            NEKKUEVRSSIDYAPLLEVDVSPENEKVLVGHVGGGWTAAVGMKFPNKISLAV
            FLNAMPDTEPNSVLYEATTPPEAKDKQFSAYGDPPTISLVCGPFISLTYH
            LSPIDHAGKLIVRPGSLIFIEDLKAKEFTBEGFGSPRVVYVIAEDKTIPTPEFRW
            MEENPVKEVKEIKGADHMPSPKDELSQLLDIAKKA"
ORIGIN
Query Match      29.0%; Score 312.4; DB 8; Length 891;
Best Local Similarity 63.4%; Pred. No. 1.5e-67;
Matches 497; Conservative 0; Mismatches 281; Indels 6; Gaps 1;
QY 31 ATAAAAATCAATGAAGGAGGAAACACTTTGTTTGTAGTACATGTCATGCCATGGAG 90
DB 37 AAAAAAAAAATGGAAGTAATGAACCATTTTGTACGCTCCACGGCTAGGCCAGGAG 96
QY 91 GTTGGAGTTGGTACAGCTAAAGCCACTGCTAGAAGCTGCAGGCCCAATGAAGTTACAGCCC 150
DB 97 CTGGGTCTATTACAGCTGAAACCTAGAAATTTGAGCGACGAGCCACCGTTGTACGGCGG 156
QY 151 TTGATTTAGCAGTTCTGGCACTGATTTGAGAAAAATAGAGAGCTTGCACACTTATG 210
DB 157 TTAATTTAGCTGTTCTGGAATCAACGAAAGAAACTCGAAGAAGTTTCGTTCAATTTG 216
QY 211 ATTATACTTTGCATTTGAGGTTGATGGAATCTCTTTTTCACAGATGAGAGGTTATAT 270
DB 217 ATTACCGCGCGGTTGTTGGAGGTTTGTGATTTCTTCTGAAATGAGAAGTAATAC 276
QY 271 TAGTGGGGCATAGTCTTGGTGTATGAATTTGGGACTTGTCTATGGAAAGATATCCACAAA 330
DB 277 TGTAGGACATAGCGCGGCTGGTATGACGGCGCTGTTGGGATGAGAAATTTCCGAATA 336
QY 331 AGATCATGCTGTTGTTTCTTGGCTGCTTTTCATGCTGATTTCTGTTTCACACTCTCTCT 390
DB 337 AGATCTCTTTAGCTGTTCTTTTGAATGCTATAATGCTGTATACCGGAAACAGACCATCTT 396
QY 391 TTGTTTTGGAACAGTATAATGAGCGGACCGGAGAGAAATGGTTGGATCTCATGTTT 450
DB 397 ATGTTTTGGAAGGTACACTGCAAGACTCCACGAGAACATGGAAGATTCCCGATTCT 456
QY 451 TACCATATGTTTCCCTGAAGAGCCACTGCATCCATGTTTGTGGCCCAAGATGTTGG 510
DB 457 CAGCCTATGAGATCCT-----CCTATTACATCTTTGGTCTGCGGCTCGATGTTCA 510
```

```
QY 511 CTCACAGCTCTACCAGCTATGCTCTCTGAGGATCTTGGCAATTAGCATCATCATTTGGTGA 570
DB 511 CATCACCTCTGTATCACCTCTCCCTTATTGAGGATCATGCAATTTGGAAAAATATTGGTAA 570
QY 571 GACCAAGCTCTTTGTTTATGGAAGACCTTATCGAAGGCCCAAGTATTTCACAGATGAACGGT 630
DB 571 GGCCAGGCTCACTATTATTAGAGGATTATTATGAAGGCCAGAAAAATTTTACAGAGGAAGAT 630
QY 631 TTGGATCAGTGAAGAGAGTTTACATTTGTGTGCACCTAGGATAAAGGCNTACCAGAGNAT 690
DB 631 TTGGATCAGTACCAAGAGTTTATGTAATAGCTGTGTAAGATAAAACATATACCACTGTAAT 690
QY 691 TCCAGCGATGCGAAATTGACAACATTTGTCCTCACTGAAGCAATAGAGATTTAAAGTGTCTG 750
DB 691 TCCAACGTTGGATGATCGAAAAACCAACCAGTTAAGAAAGTCAAGAGATCAAGGTGCTG 750
QY 751 ATCATATGCGCAATGCTATGCGAGGCCCAAAACCTTTGCGCCTCTCTCTTTGGAAATTTGCC 810
DB 751 ATCATATGCGCTATGTTCTTAACCTGATGAATATCCCAATGCTTTTGGACATAGCCA 810
QY 811 ATAA 814
DB 811 AGAA 814
RESULT 5
BT002859 935 bp mRNA linear PLN 15-JAN-2003
LOCUS Arabidopsis thaliana clone RAFL14-80-F10 (R20356) putative
DEFINITION acetone-cyanohydrin lyase (At2g23620) mRNA, complete cds.
ACCESSION BT002859
VERSION BT002859.1 GI:27754456
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
REFERENCE 1 (bases 1 to 935)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,
            Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
            Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
            Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
            Arabidopsis Full Length cDNA Clones
            Unpublished
            2 (bases 1 to 935)
            Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,
            Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
            Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
            Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
            Direct Submission
            Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'); Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M.,
            Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M., Onodera,C.S.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G.,
            Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,
            Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
```

Yamada, K. (SSP/PBEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PBEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

#### FEATURES

source  
1. .935  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosomes="2"  
/clone="RAF14-80-F10 (R20356)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."  
gene  
1. .935  
/gene="At2g23620"  
5'UTR  
1. .43  
/gene="At2g23620"  
44. 835  
/gene="At2g23620"  
/codon\_start=1  
/evidence=experimental  
/product="putative acetone-cyanohydrin lyase"  
/protein\_id="AAO22676.1"  
/db\_xref="GI:27754457"  
/translation="MSEERKQHFVLVHSGCHGAWCYKVLLEAVGHRVAVDLAA SGIDTTRSDIPTCQYSEPLTKLITSLPNDEKVLVHSGHSGGLNLAISMFKPKKI SVAVFTAPMDEHPSPFVLDFKPSNMPQEAWMGTEPEFYDSDNSGSMFSPDFMK LGIQLSPLEJLEGLMRPGSLFINLDSKMKFSDGYSVPRFIVICKEDKAIFE ERQRWIDNFPVNLVMEBETDHPMFCKPQLSDYFLKIADFV"  
misc\_difference  
341  
/gene="At2g23620"  
/note="compared to genomic sequence resulting in an amino acid sequence difference"  
/replaces="g"  
836. .935  
/gene="At2g23620"  
3'UTR  
ORIGIN

Query Match 27.4%; Score 295.2; DB 8; Length 935;  
Best Local Similarity 62.7%; Pred. No. 3.1e-63;  
Matches 494; Conservative 0; Mismatches 288; Indels 6; Gaps 2;  
QY 53 AAAACACTTTGTTTGTAGTACATGTCATGCCATGCCAGGTGGAGTTGGATGGTACAGCTAAA 112  
DB 64 ACAACATTTTGTACTAGTACATGTTCTGTCGCCATGGCGGTGGTGTGTGTGTGTGTGTAA 123  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
DB 124 GCCCGTGTAGAGCGGTGGGCCACCGCGTAATCTGTGTGACCTTGTGCTTCCGGAAT 183  
QY 173 TGA---TTTGAGAAAATAGAGGAGCTTCGCACACTTTATGATATATATCTTGGCAATGAT 229  
DB 184 AGACACAACAGAGTGTGATCACTGACATCCCAACATCGCAACAATCTCGAGGCCATTGAC 243  
QY 230 GAGATTGTAGGAATCTCTTTACGACATGAGAAGTTATATAGTGGGCGCATAGTCTGG 289  
DB 244 GAAGCTCCTGACCTCATGTCCTCAATGATGAAAAGGTTGTGCTGTTGGTGCACAGCTTTGG 303  
QY 290 TGGTATGAATTTGGGACTTGTATGAAAAGTATCCACAAAAGATCTATGCTGCTGTTT 349  
DB 304 TGGCTTGAACCTTAGCCATAGCCATGGAAGAAGTTTCCAAAANAATCTGTGCTGTATT 363  
QY 350 CTTGGCTGCTTTTCATGCTGATTCTGTTCACAACTCCTCCTTTGTTTGGAAACAGATATA 409  
DB 364 CTTGACTGCTTTTCATGCCGACACCGAACACTCACCATCTTCTGCTTGGCAAGTTGG 423  
QY 410 TGAGCGGACGCCAGCGAGAAATTGGTTGGATATCTAGTTTTCATATGATGGTTCCCTGA 469  
DB 424 AAGCAACATGCTCCTCAAGAAGCATGGATGGCGACCGAAATTCGAACCTTATGGTTCAGACAA 483

QY 470 AGAGCCACTGACATCCATGCTTTTGGGCCAAAGTTCTTGGCTCACAAGCTCTACCAGCT 529  
DB 484 TTCCGGACTGA---GTATGTTTTTTAGCCTGACTTCATGAAGTTGGGTCTCTACCAGCT 540  
QY 530 ATGCTCTCTGAGGATCTTGCATTTAGCATCATCTGTTGAGACCAAGCTCTTTGTTTAT 589  
DB 541 TTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATGAGGCCAGGATCGTTATTTAT 600  
QY 590 GGAAGACCTATTCGAAGGCCCAAGTATTTACACAGATCAACGGTTTGGATCAGTGAAGAGCT 649  
DB 601 TAAACGATTTATCGAAGATGAANAACCTTCGCGATGAGGATAGGTCGTGTTCTCTCGAT 660  
QY 650 TTACATTTGTGCTGACTGAGGATAAAGGCATACAGAGAATTTCCAGCGATGGCAATTTGA 709  
DB 661 TTTTCATAGTGTAAAGAGGACAAAGCAATTCACAGAGAACGCCAGAGATGATTTGA 720  
QY 710 CAACATTTGGTGTCTGACTGAAGCAATAGAGATTAAAGGTGCTGATCATCATGGCAATGCTATG 769  
DB 721 TAAATTTTCGGGTGAATTTAGTGATGGAGATGGAGGAGACAGATCATATGCCAATGTTCTG 780  
QY 770 CGAGCCCCCAAAACCTTTCGGCTCTCTCTTGGAAATTCGCCATAAATACAACTGATCTCT 829  
DB 781 CAAGCCTCAGCAACTCAGTGATTACTTCTGAAAATTCGGGACAAATTCGTTTAATCAAA 840  
QY 830 ACATTATG 837  
DB 841 TCTTCATG 848  
RESULT 6  
AX506625  
LOCUS AX506625 792 bp DNA linear PAT 27-SBP-2002  
DEFINITION Sequence 1320 from Patent WO0216655.  
ACCESSION AX506625  
VERSION AX506625.1 GI:23387862  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
REFERENCE  
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 1320 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG (CH)  
FEATURES  
source  
1. .792  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 27.1%; Score 292.2; DB 6; Length 792;  
Best Local Similarity 63.1%; Pred. No. 1.8e-62;  
Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;  
QY 53 AAAACACTTTGTTTGTAGTACATGTTGCATGCCATGCCAGGTGGAGTTGGTACAGCTAAA 112  
DB 21 ACAACATTTTGTACTAGTACATGTTCTGTCGCCATGGCGGTGGTGTGTGTGTGTGTAA 80  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGTTACAGCCCTTGTATGATGATATCTTTGCCATTGAT 172  
DB 81 GCCGCTGTAGAGCGGTGGGCCACCGCGTAATCTGTGGAATTTAGCTGCTCCGGAAT 140  
QY 173 TGA---TTTGAGAAAATAGAGGAGCTTCGCACACTTTTATGATATATCTTTGCCATTGAT 229  
DB 141 AGACACAACAGAGTGTGATCACTGACATCCCAACATCGCAACAATCTCGAGGCCATTGAC 200  
QY 230 GGAGTTGATGGAATCTCTTTACGACAGATGAGAAGTTATATTTAGTGGGCGCATAGTCTTGG 289

Db 201 GAAGCTCCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTCGTTGGTCACAGCTTTGG 260  
Qy 290 TGGTATGAATTTGGGACTTTGCTATATGGAAGATATCCAAAAGATCTATGCTGCTGTTTT 349  
Db 261 TGGCTTGAATTTAGCCATAGCCATGGAAGTTTCCGAAAAAATCTCTGTGCTGTATT 320  
Qy 350 CTTGGCTGCTTTTCATGCTGATTCGTTTCAACAACCTCTCTTTGTTTGGACAGATATA 409  
Db 321 CTTGACTGCTTTTCATGCTGCGGACACCGAACCACTCACCATCTCTTCGCTTGGCAAGTTTGG 380  
Qy 410 TGAGCGGACGCCAGCGGAAATTTGGTTGGATATCTAGTTTACCATATGTTTCCCTTGA 469  
Db 381 AAGCAACATGCTCAGAGACATGATGGGACCGAATTCGAACCTTATGTTTCAGACNA 440  
Qy 470 AGAGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTTGGCTCAACAGCTCTACAGCT 529  
Db 441 TTCCGGACTGA--GTATGTTTTTAGCCCTGACTTCATGAAGTTGGGTCTCTACAGCT 497  
Qy 530 ATGCTCTCCTGAGGATCTTGCATTAGCATCATCATTTGGTGAGACCAAGCTCTTTCTTTAT 589  
Db 498 TTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGGCCAGGATCGTTATTTAT 557  
Qy 590 GGAAGACCTTATCGAAGGCCAAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGT 649  
Db 558 TAACGATTTATCGAAGATGAAGAACTTCTCGGATGAAGATATGGTCTGTTCTCGAGT 617  
Qy 650 TTACATTTGTGCTGAGGATTAAGGCATACAGAAAGTATTCAGAGCGATGGCAATTTGA 709  
Db 618 TTTTCAATGTGTAAAGAGGACAAAGCAATTCAGAAAGAACCCAGAGATGGATTTGA 677  
Qy 710 CAACATTTGTGCTGACATGAGCAATAGAGATTAAGGTGCTGATCATCGCAATGCTATG 769  
Db 678 TAAATTTCCGGTGAATTTAGTGAATGAGATGAGGAGACAGATCATATGCCAATGTTCTG 737  
Qy 770 CGAGCCCCAAAACCTTTGGCCCTCTCTCTTGGAAATGGCCATAAATAC 818  
Db 738 CAAGCCTGAGCACTCAGTGATTAATCTCTGAAATCGCGGACAAATTC 786

RESULT 7  
LOCUS BT014881 792 bp mRNA linear PLN 05-JUN-2004  
DEFINITION Arabidopsis thaliana At2g23620 gene, complete cds.  
ACCESSION BT014881  
VERSION BT014881.1 GI:48310670  
KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 792)  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.  
JOURNAL Arabidopsis ORF clones  
TITLE Unpublished  
REFERENCE 2 (bases 1 to 792)  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.  
JOURNAL Direct Submission  
TITLE Submitted (05-JUN-2004) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

FEATURES  
source Location/Qualifiers  
1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U67545"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
1..792  
/note="unknown protein"

CDS

/codon\_start=1  
/product="At2g23620"  
/protein\_id="AA141864.1"  
/db\_xref="GI:48310671"  
translation="MSEKRSQHEFLVHSGCHGAWCYKVKPLLEAVGHRVTAVDLAA  
SGIDTTSITDIPTEQVSEPLTKLTLSPNDEKVLVLVGHSGFLGSLNLAIAEMKPEPKI  
SVAVPLTAFMPDTEHSPFLDKFGSNMPOEAMNGTBPFPYGSNLSGSMFSPDFMK  
LGLYLQSPVEDLELGLLLMPPSLPINDLSKMKNPSDEGYSVPRVFTVCKEDRAIPE  
ERQRWMDINFPVNLVMEETDHPMFCKPQOLSDFLKIADKFV"

ORIGIN

Query Match 27.1%; Score 292.2; DB 8; Length 792;  
Best Local Similarity 63.1%; Pred. No. 1.8e-62;  
Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;  
Qy 53 AAAACACTTTCTTTTGTACATGTCATGCGCATGCGAGTTGGAGTTGGTACAAGCTAAA 112  
Db 21 ACNACATTTTGTACTAGTACATGTTTCGTGCATGCGCGCTGCTGTTCAAGGTTAA 80  
Qy 113 GCCACTGCTAGACGTCGAGGCCATAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
Db 81 GCCGCTGCTAGAGCGGTGGGCCACCGCGTAACCTGCTGTGGACTTAGCTGCCGAAT 140  
Qy 173 TGA---TTTGGAAAAATAGAGAGCTTTCGCACACCTTTATGATTATATCTTTGCCATTGAT 229  
Db 141 AGACACACGAGGTGCTGATCACTGACATCCACATCGCAACAATACTCGGAGCCATTGAC 200  
Qy 230 GGAGTTGATGAATCTCTTTTACAGCAGATGAGAGGTTATATAGTGGGGCATAGTCTTGG 289  
Db 201 GAACTCCTGACCTCATTTGCCAAATGATGAAAAGGTTGTGCTCGTTGCTGACAGCTTTGG 260  
Qy 290 TGGTATGAATTTGGGACTTGTCTATGAAAAGTATCCACAAAAGATCTATGCTGCTGTTTT 349  
Db 261 TGGCTTGAACCTTAGCCATAGCCATGAAAAAGTTTCCCGAAAAAATCTCTGCTGCTGTATT 320  
Qy 350 CTTGCTGCTTTTCATGCTGATTTCTGTTCACAACTCTCTCTCTTTTGGAAACAGTATAA 409  
Db 321 CTTGACTGCTTTTCATGCTGCGGACACCGAACACTCACCATCTCTGCTTTGGACAGTTGG 380  
Qy 410 TGAGCGGACGCCAGCGGAAATTTGGTTGGATCTCAGTCTTTTACCATATGTTTCCCTTGA 469  
Db 381 AAGCAACATGCTCAGAGACATGATGGGACCGAATTCGAACCTTATGTTTGGTTCAGACAA 440  
Qy 470 AGAGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTTGGCTCAACAGCTCTACAGCT 529  
Db 441 TTCCGGACTGA--GTATGTTTTTAGCCCTGACTTCATGAAGTTGGGTCTCTACAGCT 497  
Qy 530 ATGCTCTCCTGAGGATCTTGCATTAGCATCATCTATGTTGGTGAGACCAAGCTCTTTGTTAT 589  
Db 498 TTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGGCCAGGATCGTTATTTAT 557  
Qy 590 GGAAGACCTTATCGAAGGCCAAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGT 649  
Db 558 TAAACGATTTATCGAAGATGAAAAAATCTTCTCGGATGAAGGATATGGGTCTGTTCTCGAGT 617  
Qy 650 TTACATTTGTGCTGACTGAGGATTAAGGCATACCGAAGAAATTCACAGCATGCAAAATGA 709  
Db 618 TTTTCAATGTGTAAAGAGGACAAAGCAATTCAGAAAGAACCCAGAGATGGATTTGA 677  
Qy 710 CAACATTTGTGCTGACTGAAGCAATAGAGTAAAGGTGCTGATCATCGCAATGCTATG 769  
Db 678 TAAATTTCCGGTGAATTTAGTGAATGAGATGAGGAGACAGATCATATGCCAATGTTCTG 737  
Qy 770 CGAGCCCCAAAACCTTTGGCCCTCTCTCTTGGAAATTCGCCATAAATAC 818  
Db 738 CAAGCCTCAGCAACTCAGTGATTAATCTCTGAAATTCGCGGACAAATTC 786

RESULT 8  
LOCUS AY086590  
DEFINITION Arabidopsis thaliana clone 26006 mRNA, complete sequence.  
ACCESSION AY086590

AY086590 870 bp mRNA linear PLN 14-APR-2003  
Arabidopsis thaliana clone 26006 mRNA, complete sequence.  
AY086590

AY086590.1 GI:21405300  
FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 870)  
Haas, B.J., Volfovsky, N., Town, C.D., Troukhan, M., Alexandrov, N.,  
Feidmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.  
Full-length messenger RNA sequences greatly improve genome  
annotation  
Genome Biol. 3 (6), RESEARCH0029 (2002)  
22089475  
12093376  
2 (bases 1 to 870)  
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feidmann, K.  
Full-length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 870)  
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feidmann, K.  
Direct Submission  
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA  
This clone sequence is one of 5,000 Ceres full-length cDNAs made  
available to TIGR and Genbank. The following quality assessment of  
this set was done by comparison with known proteins: two percent of  
the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have  
frame shifts in a coding region. A sequence is considered to be  
5'-truncated if it lacks the translation initiation start (ATG). A  
c-terminal end of the encoded protein. Please note that these cDNA  
sequences are derived from the Ws or Laer ecotypes and therefore  
may contain polymorphisms when compared to sequences from Col-0.  
Genet carried out the library production and sequencing of the  
full-length clones. Ceres, Inc. carried out the clustering of the  
5' sequences, selection of clones, and sequence assembly.  
Location/Qualifiers  
1..870  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="26006"  
21..812  
/codon\_start=1  
/product="putative acetone-cyanohydrin lyase"  
/protein\_id="AA063450.1"  
/db\_xref="GI:21554666"  
/translations="MSERRKQHFVLVHGAGCWYKVPFLLEALGHRVTDLAA  
SGDITFRSITDTCQYSEPLQMLTSLPNDEKVLVGHSGLSLALAMDKFDKI  
SVSFVTFAMPDPKSPFVEKFPASSMTPEGMWSELETGYGDSNLSGLVFFSTPMK  
HRLYQSPVEDLELGLLLKPSLSLFINELSKWENFSEKGVSPRAYIVCKEDNIIE  
DQRWMIHNPANLVIEETDHPWFCKPQVLSHLALADNFS"  
ORIGIN  
Query Match 25.3%; Score 273; DB 8; Length 870;  
Best Local Similarity 60.3%; Pred. No. 1-2e-57;  
Matches 488; Conservative 0; Mismatches 315; Indels 6; Gaps 2;  
QY 13 AAAAGAACTACAAAGCATATAATTCAAATGAAGGAAGGAAACACTTTGTTTATGATAC 72  
|||||  
Db 1 AAAAAGTACGAAAGGAAAATATGAGTGAAGGAAGGAAACAACTTCGTCTAGTAC 60  
|||||  
QY 73 ATGGTCATGCCATGGAGGTGGATGGTGGTACAACTGAAGCCACTGAGAGCTGCAG 132  
|||||  
Db 61 ATGGTCGTCGCCACGGCGCATGGTGGTGGTACAAAGGTAAAGCTCTCTCGAGGCTTGG 120  
|||||  
QY 133 GCCATAAGGTTACAGCCCTTGATTAGCAGCTTCTGCACTGA---TTTGAGAAAAATAG 189

Db 121 GCCATCGTGTAAACCGCCTTAGACCTAGCTGCTTCGGTATAGACACAACAGGTCAATCA 180  
|||||  
QY 190 AGGAGCTTCGCACACTTTATGATATATCTTCGCCATTGATGGAGTTGATGGAATCTCTTT 249  
|||||  
Db 181 CTGCAATTTCTACATGTGTGAACAATATTTGAGCCATTGATGAGCTAAATGATCTTATGC 240  
|||||  
QY 250 CAGCAGATGAGAAGGTTATATTAGTGGGCGCATAGTCTTCGGTGGTATGAAATTTGGGACTTGG 309  
|||||  
Db 241 CGAATGATGAGAAGGTTGTACTCGTGGTTCATAGCTTTGAGAGTTTGAGTTAGCCTTAG 300  
|||||  
QY 310 CTATGGAAGATATCCAAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCGCTG 369  
|||||  
Db 301 CCATGGATAAAGTTTCCCGATAAAATCTCTGTCTCTGTCTTCGTGACTGCAATTCATGCCCG 360  
|||||  
QY 370 ATTCTGTTCCAAACTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGAGCGCCGCGAGA 429  
|||||  
Db 361 ACACAAACACTCACCATCGTTCGTCGAGGAAAGTTTGCAGAGCAGCATGACACCAAGAG 420  
|||||  
QY 430 ATTGGTTGGATACTCAGTCTTTTACCATATGTTTCCCTCGAAGAGCCACTGACATCCATGT 489  
|||||  
Db 421 GATGGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATAATTCCGGCTTGTCTGTGT 477  
|||||  
QY 490 TTTTGGGCCAAAGTTCTTGGCTCAAGCTCTACAGCTATGCTCTCTCTGAGGATCTTGG 549  
|||||  
Db 478 TCTTCAGCACCGACTTCATGAAGCACCGTCTCTACCAACTTTCTCTGTGGAGGATCTTGG 537  
|||||  
QY 550 CATTAGCATCATCTGGTGAGACCAAGCTCTTTGTTATGGAAGACCTTCCGAGGCCA 609  
|||||  
Db 538 AGCTTGGATGTGCTTCTAAAGAGGCGCTAGTGTCAATGTTTAAATGAATTTGGAAGATGG 597  
|||||  
QY 610 AGTATTTCACAGATCAACGGTTTGGATCAGTCAGTGAAGAGAGTTTACATTGTGTGCACGTGAGG 669  
|||||  
Db 598 AGAATTTCTGAGAAAGGTTAGTGTGTTCTTCTCGAGCTTACATTGTGTGCAAGAGG 657  
|||||  
QY 670 ATAAAGGCATACAGAAGAAATTCAGCGATGCGCAAATTGACAACTATGTTGTCTACTGAAG 729  
|||||  
Db 658 ACAACATATCTCGGAAGACCATCAACGATGGATGATCCATAATTTATCCGGGGAATTTAG 717  
|||||  
QY 730 CAATAGAGTTAAAGTGTCTGATCAGTCGCAATGCGCAATGCTATGCGAGCCCAAAAACCTTTGCG 789  
|||||  
Db 718 TCATTGAGATGAAGAGAGCGGATCATATGCCAATGTTTTCACAAACCTCAAGTACTAAGTG 777  
|||||  
QY 790 CCTCTCTCTTGGAAATTTGCCATAAATAC 818  
|||||  
Db 778 ACAATCTATTGGCAATCGCTGACAATTC 806  
|||||  
RESULT 9  
AF361627  
LOCUS AF361627  
DEFINITION Arabidopsis thaliana At2g23600/F2686.25 mRNA, complete cds.  
ACCESSION AF361627.1 GI:13605602  
VERSION AF361627.1  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 932)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,  
Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M.,  
Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Phan, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinozaki, K.,  
Davis, R.W., Theologis, A. and Ecker, J.R.  
Arabidopsis cDNA clones  
Unpublished  
2 (bases 1 to 932)  
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,  
Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,





```
QY 115 CACTGTAGAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174
Db 83 CTCTTCTCGAGGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTTCGGGTATAG 142
QY 175 A---TTTGAGAAAATAGAGAGCTTCGCACACTTATGATTTATATCTTTGCGCATTTGATGG 231
Db 143 ACACAACCGGTCAATCACTGACATTTCTACATGTGTGAACAATATCTGAGCCCAATGATGC 202
QY 232 AGTTGATGGAACTCTTTTCAGCAGATGAGAAGGTTATATTAGTGGGCGCATAGTCTTGGTG 291
Db 203 AGCTAATGACTTCAATGCCCAATGATGAGAAGGTTGTACTCGTGTGTCATAGCTTTGGAG 262
QY 292 GTATGAATTTGGAGCTTGTATGAAAAGTATCCAAAAGATCTATGCTGCTGTTTCT 351
Db 263 GTTTGAGTTTGGCTTAGCCATGGATAAGTTTCCGATAAAATCTCTGCTCTGCTCTCG 322
QY 352 TGCTGCTTTTCAGCCTGATTTCTGTTACAACTCTCTCTTTGTTTGGACAGTAAATG 411
Db 323 TGACTGCAATTCATGCCCGCACCAAAACACTTCACCATCGTTCTGTCGAGGAAAGTTTGCAA 382
QY 412 AGCGGAGCGCCAGCGCAGAAATTTGGTGGATCTCAGTATTTTACCATATGGTTCCCTGGAAG 471
Db 383 GCAGCATGACACAGAGGATGATGGCTCTGAGCTCGAGACATATGTT---CAGATA 439
QY 472 AGCCACTGACATCCATGTTTTTTGGCCAAAGTTCTTGGCTCAAGCTCTACCACTAT 531
Db 440 ATTCGCGCTTGTCTGTGTTCTTCAGCAGCAGCTTCATGAAGCACCGCTCTCTACCACTTT 499
QY 532 GCTCTCTGAGGATCTTGCAATGATGATCATCTGTTGGAGACCAAGCTCTTTGTTATGG 591
Db 500 CTCCTGTGGAGGATCTTGAGCTTGGATGCTTCTTAAAGAGCGCTAGTTCATTGTTTATTA 559
QY 592 AAGACCTATCGAAGGCCAAGTATTTCCAGATGACGCTTTTGGATCAGTCAAGAGAGTTT 651
Db 560 ATGAATATCGAAGATGAGAACTTTTCTGAGAAAGGATGATGATCTGTTCTCGAGCTT 619
QY 652 ACATTGTTGCTCACTGAGGATAAAGGATACCAAGAAATTTCCAGCGATGGCAATTGACA 711
Db 620 ACATTGTTGCAAGAGGACAACTATTCGGAAGACCATCAACGATGATGATCCATA 679
QY 712 ACATTGTTGCTCACTGAGCAATGAGATTAAGGTTGCTGATTCATGCAATGCAATGCTATGG 771
Db 680 ATTTATCGCGGAAATTTAGTGAATGAGATGAGAGACTGATCATATGSCCAATGTTTGTGA 739
QY 772 AGCCCCAAAACCTTTGCGCTCTCTCTTGGAAATTCGCCATAAATAC 818
Db 740 AACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGACAAATTC 786
```

## RESULT 11

AY056068 792 bp mRNA linear PLN 30-SEP-2001  
Arabidopsis thaliana At2g23600/F26B6.25 mRNA, complete cds.

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Koesema, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P.,  
Ban, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,  
Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,  
Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,  
Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,  
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## TITLE

Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

## COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Koesema, E., Chen, H.,  
Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P., Ban, J., Bowser, L.,  
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,  
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,  
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
Davis, R.W., Theologis, A., and Ecker, J.R.

Koesema, E. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
contributed equally to this work as PIs.

## FEATURES

## source

1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U12118"  
/ecotype="Columbia"  
/note="this clone is in PUNI 51"

## CDS

1..792  
/note="putative acetone-cyanohydrin lyase"  
/codon\_start=1  
/product="At2g23600/F26B6.25"  
/protein\_id="RAL06968.1"  
/db\_xref="GI:15810085"  
/translation="MSEERKQHFVLVHGCHGAWCYKVKPILLEALGHRVTALDLAA  
SGIDTTSITDISTCEQYSEFLMQLMTPNDKVLVGHGSLSLALAMDKFPDKI  
SVSVFTAFMPDTHKSPFVEKFASSMTPEGMWSELETGSDNSGLSVFSTDFMK  
HRLYOLSPVEDLELGLLKRPSLPINELSKMENPEKGYGVPRAIVCKEDNIIE  
DQRWMIHNPANLVIEETDHPMFCKPQLSLDHLAIADNFC"

## ORIGIN

Query Match 25.1%; Score 271; DB 8; Length 792;  
Best Local Similarity 61.4%; Pred. No. 3.7e-57;  
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;  
QY 55 AACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGGTGGATGGTGTACAGCTAAAGC 114  
Db 23 AACACTTCTGTGTAGTACATGGTGCATGCCATGGAGGTGGATGGTGTACAGCTAAAGC 82  
QY 115 CACTCTAGAGCTGCAGGCCATAAGTTCACAGCCCTTGATTTAGCAGCTTCTGGCAGCTG 174  
Db 83 CTCTTCTCGAGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTTCGGGTATAG 142  
QY 175 A---TTTGAGAAAATAGAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTGATGG 231  
Db 143 ACACAACCGAGTCAATCACTGACATTTCTACATGTGAACAATATTTCTGAGCCATTGATGC 202  
QY 232 AGTTGATGGAATCTCTTTTCAGCAGATGAGAAGGTTATATTAGTGGGCGCATAGTCTTGGTG 291  
Db 203 AGCTAATGACTTCAATTCGCGCAATGATGAGAAGGTTGTACTCGTGGTGCATAGCTTTGGAG 262



Matches		472;	Conservative	0;	Mismatches	295;	Indels	6;	Gaps	2;
Qy	55	AACACTTTGTTTGTAGTACATGTCATGTCATGAGGTGGAGTTGGTGTGTAACAAGCTAAAGC 114								
Db	23	AACACGTCGTTCTAGTACATGTCGTTGCCATGCGCTGGTCTGGTACAAGTTAAGC 82								
Qy	115	CACTGCTAGAGCTGCAGGCATTAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174								
Db	83	CGCAGCTCGAGGCTTCTGGCCACCGGTAACCGCGTAGATCTAGTGCCTCGGTATAG 142								
Qy	175	ATTG---AGAAAAATAGAGAGCTTCGCACACTTTATGATTAATCTTTGGCACTATG 231								
Db	143	ACATGACCGAGTCAATACAGATATATCCACATGCGAACTACTCAGAGCCATTGATG 202								
Qy	232	AGTTGATGGAATCTCTTTAGTACATGTCATGCCATGAGAGTTGGAGTTGGTACAAGCTAAAGC 114								
Db	203	AGCTAATGACCTCACTACAGATGATGAGAAGTTGTGCTTCTGTCATAGCTTAGGAG 262								
Qy	292	GTATGAATTTGGGACTTGTCTATGGAAGAGTATCCAAAGATCTATGCTGCTTTCT 351								
Db	263	GTTTGAGTTTGTAGTATGCGCATGATGTTTCCGACCAAAATCTCTGTTCTGCTTTG 322								
Qy	352	TGCTGCTTTCATGCTGCTGATCTGTTCAACCTCTCTTTGTTTGGAAACAGTAAATG 411								
Db	323	TGACTGCTATGATGCCAGACCAACACTTCACCTCTTCGTATGGGATAAGCTAAGAA 382								
Qy	412	AGCGGACCGCAGCGAGAAATGTTGATGATCACTCAGTTTATACCATATGTTCCCTGAAG 471								
Db	383	AGAAACTTCACGAGGAGGATGTTAGACACCGTGT---ACGAGCGAAGAACCTGATT 439								
Qy	472	AGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTGCTCACAAGCTCTACCAAGCTAT 531								
Db	440	TTCTTAGCGAGTTTGGATTTTGGACCAAGTTCATGCCAAGAACTTGTATCATGTTGT 499								
Qy	532	GCTCTCTGAGGATCTTGCAATTTAGCATCATCTGTTGAGCAAGCTCTTTGTTTATGG 591								
Db	500	CTCCAGTCCAAGATCTGAATTTGGGAAAATGTTGGTGGGGCAAAACCCATTGATTAGA 559								
Qy	592	AAGACTTATCGAAGCGCAAGTATTTTACAGATGAACCGTTTGGATCAGTGAAGAGAGTTT 651								
Db	560	AAGATATGCCAGAGAGAGAGCTTCAGTGAAGAGGATACGGATCCGTTACACGTATAT 619								
Qy	652	ACATTGTGTGCACTGAGGATAAGGCATACCAAGAGAATTCAGGATGGCAAAATGACA 711								

RESULT 14	
BT015031	
LOCUS	792 bp mRNA linear PLN 10-JUL-2004
DEFINITION	Arabidopsis thaliana At2g23610 mRNA, complete cds.
ACCESSION	BT015031
VERSION	BT015031.1
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1 (bases 1 to 792)
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.
TITLE	Arabidopsis ORF clones
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 792)
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.

TITLE		Direct Submission
JOURNAL		Submitted (10-JUL-2004) Salk Institute Genomic Analysis Laboratory (SIGNAL), plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES		Location/Qualifiers
source		1..792
		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
		/chromosomes="2"
		/clone="U67063"
		/ecotype="Columbia"
		/note="This clone is in PUN1 51"
CDS		1..792
		/note="unknown protein"
		/codon_start=1
		/product="At2g23610"
		/protein_id="AA170482.1"
		/db_xref="GI:50198958"
		/translation="MSBERKOHVVLVHGACHGAWCYKVKPOLBASGHRVTVADLAA
		SGIDMTRSIIDISTCEQYSEPLQMLMTSLPDDKVLVGHSLGSLAMAMDMFTKI
		SVSVFTAMMPTKSPSPFVNDKLKRETSREBLDTVTFSEKDPSPSEFWIPGPFMA
		KNLYQLSPVDLELAKMLVRANPLIKDMAERRSPSEBGSVTRIFVCGKDLVSPSE
		DYQSRSMISNFPPEYMEIKDADHMPFSPKPOLCALLLEIANKYA"
ORIGIN		
Query Match		24.9%; Score 269; DB 8; Length 792;
Best Local Similarity		61.1%; Pred. No. 1.2e-56;
Matches		472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
Qy	55	AACACTTTGTTTGTAGTACATGTCATGCCATGAGAGTTGGAGTTGGTACAAGCTAAAGC 114
Db	23	AACACGTCGTTCTAGTACATGTCGTTGCCATGCGCTGGTCTGGTACAAGTTAAGC 82
Qy	115	CACTGCTAGAGCTGCAGGCATTAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174
Db	83	CGCAGCTCGAGGCTTCTGGCCACCGGTAACCGCGTAGATCTAGTGCCTCGGTATAG 142
Qy	175	ATTG---AGAAAAATAGAGAGCTTCGCACACTTTATGATTAATCTTTGCCATGATG 231
Db	143	ACATGACCGAGTCAATACAGATATATCCACATGCGAACTACTCAGAGCCATTGATG 202
Qy	232	AGTTGATGGAATCTCTTTAGTACATGTCATGCCATGAGAGTTATATTAGTGGGCACTAGTCTTTGGTG 291
Db	203	AGCTAATGACCTCACTACAGATGATGAGAAGTTGTGCTTCTGTCATAGCTTAGGAG 262
Qy	292	GTATGAATTTGGGACTTGTCTATGGAAGAGTATCCAAAGATCTATGCTGCTTTCT 351
Db	263	GTTTGAGTTTGTAGTATGCGCATGATGTTTCCGACCAAAATCTCTGTTCTGCTTTG 322
Qy	352	TGCTGCTTTCATGCTGCTGATCTGTTCAACCTCTCTTTGTTTGGAAACAGTAAATG 411
Db	323	TGACTGCTATGATGCCAGACCAACACTTCACCTCTTCGTATGGGATAAGCTAAGAA 382
Qy	412	AGCGGACCGCAGCGAGAAATGTTGATGATCACTCAGTTTATACCATATGTTCCCTGAAG 471
Db	383	AGAAACTTCACGAGGAGGATGTTAGACACCGTGT---ACGAGCGAAGAACCTGATT 439
Qy	472	AGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTGCTCACAAGCTCTACCAAGCTAT 531
Db	440	TTCTTAGCGAGTTTGGATTTTGGACCAAGTTCATGCCAAGAACTTGTATCATGTTGT 499
Qy	532	GCTCTCTGAGGATCTTGCAATTTAGCATCATCTGTTGAGCAAGCTCTTTGTTTATGG 591
Db	500	CTCCAGTCCAAGATCTGAATTTGGCGAAAATGTTGTTGGGGCAAAACCCATTGATTAGA 559
Qy	592	AAGACTTATCGAAGCGCAAGTATTTTACAGATGAACCGTTTGGATCAGTGAAGAGAGTTT 651
Db	560	AAGATATGCCAGAGAGAGAGCTTCAGTGAAGAGGATACGGATCCGTTACACGTATAT 619
Qy	652	ACATTGTGTGCACTGAGGATAAAGGCATACCAAGAAATTTCCAGCGATGCGAAATTGACA 711
Db	620	TTATTGTATGCGAAGAGATCTTGTGTCACCCGAGATTACCAAGCATCGATCAGCA 679
Qy	712	ACATTGTTGTCATGAAGCAATAGAGATTAAGAGTGTGATCATCATGCGCAATGCTATGCG 771
Db	680	ACTTTCCCCAAAGAAGTAATGGAGATCAAGACGCGAGATCATATGCCAATGTTCTCCA 739
Qy	772	AGCCCCAAAACCTTTGCGCTCTCTCTTGGAAATTCGCCATAAATACAACCTGA 824
Db	740	AGCCTCAACACTATGTGCTCTCTCTTGGAGATTGCAATAAATATGCTTAA 792

```
Db      620 TTATTGTATCGGAAGGATCTTGTCACCCGAGATTACCAGCGATCGATGATCAGCA 679
Qy      712 ACATTGGTGTACTAGCAATAGAGATTAAAGGTGCTGATCACATGGAATGCTATGCG 771
Db      680 ACTTTCCCAAGAAAGTAATGGAGATCAAGAGCGCAGATCATATGCGCAATGTTCTCCA 739
Qy      772 AGCCCCAAAATTTTGGCGCTCTCTCTGGAAATTTGCCCAATAAATACAACATGA 824
Db      740 AGCCTCAACAACATATGTGCTCTTCTTGGAGATTGCAAAATAAATATGCTTAA 792

RESULT 15
LOCUS   BT014901
DEFINITION Arabidopsis thaliana At2g23610 gene, complete cds.
ACCESSION BT014901
VERSION   BT014901.1 GI:48596978
KEYWORDS FLI CDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 878)
          Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
          Arabidopsis cDNA clones
          Unpublished
REFERENCE 2 (bases 1 to 878)
          Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
          Direct Submission
          Submitted (11-JUN-2004) Salk Institute Genomic Analysis Laboratory
          (SIGNAL), Plant Biology Laboratory, The Salk Institute for
          Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
          USA
FEATURES             Location/Qualifiers
     source            1..878
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="S67063"
                     /ecotype="Columbia"
                     /note="this clone is in pUNI 51"
     CDS               8..799
                     /note="unknown protein"
                     /codon_start=1
                     /product="At2g23610"
                     /protein_id="AAT4603.1"
                     /db_xref="GI:48596979"
                     /translation="MSEERKQHVVLVHGACHGAWCWYKVKPQLRSGHRVTAVDLAA
                     SGIDMTSRITDSTCEQYSEPLMQLSLPDDERKVLVGHSLGSLAMVDMFPFKI
                     SVSFVTAMPTKHSFVWDLKRKTSREWLDTFTSEKPDFSFBFWIFGPFFMA
                     KNYQLSPVQDLLEKMLVRANPLIKDMAERRSFSEGYGSVTRIFIVCGKDLVSPE
                     DYQRSMISNFPPEVMEIKDHPMFSPQQLCALLLEIANKYA"

ORIGIN
Query Match      24.9%; Score 269; DB 8; Length 878;
Best Local Similarity 61.1%; Pred. No. 1.2e-56;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
Qy      55 AACACTTTGTTTAGTACATGGTGCATGCCATGGAGGTTGGAGTTGGTACAAGCTAAAGC 114
Db      30 AACACGCTCGTTCTAGTACATGGTGTGCTTGGCATGGCGCCCTGGTGTGATACAAGGTTAAGC 89
Qy      115 CACTGTCTAGAGCTGCAGGCGCATAAAGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTG 174
Db      90 CGCAGCTCGAGGCTTCTGGCCACCGCGTAACCGCCGTAGATCTAGCTGGCTCGGTATAG 149
Qy      175 ATTTC---AGAAAAATAGAGGAGCTTCGCACACTTTATGATTATATCTTTGCCATTGATGG 231
Db      150 ACATGACCAAGTCAATCACAGATATATCCACATGCGAACAAATACTCAGAGCCATTGATGC 209
Qy      232 AGTTGATGGATCTCTTTAGCGAGATGAGAAGGTTATATTAGTGGGGCATAGTCTTGGTG 291
```

```
Db      210 AGCTAATGACCTCACTACCAGATGATCAGAAGGTTGTGCTTGTGGTCATAGCTTAGGAG 269
Qy      292 GTATGAATTTGGGACTCTTGCTATATGAAAAGTATCCAAAAAGATCTATGCTGCTGTTTCT 351
Db      270 GTTTGAGTTTAGCTATGGCCATGATATGTTTCCGACCAAAATCTCTGTTCTGCTTTG 329
Qy      352 TGGCTGCTTTCATGCCCTGATTTCTGTTCAACTCTCTCTTGTGTTTGGACAGTATATG 411
Db      330 TGACTGTCTATGATGCCAGACACCAACACTCACCATCTCTGATATGGATTAAGCTAAGAA 389
Qy      412 AGCGGAGCCAGCCGAGAAATTTGGTATGATCTCAGTTTATTTTACCATATATGTTTCCCTCAAG 471
Db      390 AAGAAACTTTCAGAGAGAGATGTTAGACACCGTGT---ACGAGCGAGAAACCTGATT 446
Qy      472 AGCCACTGACATCCATGTTTTTTGGCCCAAAAGTTCTTTGGCTCACAAAGCTCTACAGCTAT 531
Db      447 TTCCTAGCGAGTTTGGATTTTGGACCAAGAAATTCATGGCCCAAGAACTTGTATCAGTTGT 506
Qy      532 GCTCTCCTGAGGATCTTGCATTAGCATCATCTTGGTGAACCGTTTGGATCAGTGAAGAGAGTTT 591
Db      507 CTCCAGTCCAAAGATCTTGAATTTGGCGAAATGTTGGTGAGGGCAACCCCATTTGATTAGA 566
Qy      592 AAGACCTATCCGAAGCCCAAGTATTTTCAAGATGAACCGTTTGGATCAGTGAAGAGAGTTT 651
Db      567 AAGATATGGCAGAGAGAGAAGCTTTCAGTGAGGAAGGATACGGATCCGTTACACGTATAT 626
Qy      652 ACATTGTGTGCACTGAGGATAAAGCATACAGAAAGAAATTCACGGATGGCAAAATTGACA 711
Db      627 TTATTGTATGGGAAAAGGATCTTGTGTCAACCGAAGATTACCGCGATCGATGATCAGCA 686
Qy      712 ACATTGGTGTCACTGAAAGCAATAGAGTTAAAGGTGCTGATGATCAGATGSCAATGCTATCG 771
Db      687 ACTTTCCCCCAAAAGAGTAATGGAGATCAAAAGACGCAGATCATATGCCAATGTTCTCCA 746
Qy      772 AGCCCCAAAAACTTTTGGCGCTCTCTTTGGAAATTTGCCCATATAAATACAACCTGA 824
Db      747 AGCCTCAACAACATATGTGCTCTTCTTTGGAGATTGCAAAATAAATATATGCTTAA 799
```

Search completed: September 12, 2005, 04:45:50  
Job time : 4689.27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 11, 2005, 22:50:50 ; Search time 591.255 Seconds  
(without alignments)  
10803.121 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaagaagaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1079	100.0	1079	8 ADA26498	Ada26498 Tobacco s
2	295.2	27.4	918	8 ADA26530	Ada26530 Arabidops
3	292.2	27.1	792	6 ABZ13515	Abz13515 Arabidops
4	292.2	27.1	816	3 AAC42344	Aac42344 Arabidops
5	276.2	25.6	950	8 ADA26531	Ada26531 Arabidops
6	273	25.3	870	3 AAC34411	Aac34411 Arabidops
7	272.6	25.3	864	3 AAC47519	Aac47519 Arabidops
8	271.6	25.2	1152	8 ADA26532	Ada26532 Arabidops
9	271	25.1	792	6 ABZ13746	Abz13746 Arabidops
10	269	24.9	792	3 AAC43133	Aac43133 Arabidops
11	269	24.9	792	6 ABZ12412	Abz12412 Arabidops
12	269	24.9	792	12 ADN74620	Adn74620 Thale cre
13	256	23.7	1109	8 ADA26533	Ada26533 Arabidops
14	252.6	23.4	1046	8 ADA26535	Ada26535 Arabidops
15	243.2	22.5	928	8 ADA26538	Ada26538 Arabidops
16	242.8	22.5	823	3 AAC42529	Aac42529 Arabidops
17	240.8	22.3	834	8 ADA26537	Ada26537 Arabidops
18	239.8	22.2	771	6 ABZ13473	Abz13473 Arabidops
19	239.4	22.2	770	8 ADA67827	Ada67827 Arabidops
20	234.2	21.7	1056	3 AAC47895	Aac47895 Arabidops

21	232.6	21.6	909	3 AAC47800	Aac47800 Arabidops
22	232.6	21.6	984	8 ADA26534	Ada26534 Arabidops
23	224.8	20.8	783	8 ADA26536	Ada26536 Arabidops
24	215	19.9	903	3 AAC34141	Aac34141 Arabidops
25	215	19.9	908	8 ADA26539	Ada26539 Arabidops
26	212.8	19.7	828	3 AAC41733	Aac41733 Arabidops
27	186	17.2	1091	2 AAT36351	Aat36351 Hevea bra
28	184.2	17.1	817	2 AAT86824	Aat86824 Hevea bra
29	169	15.7	777	12 ADQ16305	Adq16305 Nucleotid
30	168	15.6	795	8 ADA70991	Ada70991 Rice gene
31	156.6	14.5	777	12 ADQ16306	Adq16306 Nucleotid
32	149	13.8	422	3 AAC43356	Aac43356 Arabidops
33	145.8	13.5	1325	3 AAC47916	Aac47916 Arabidops
34	143.4	13.3	984	3 AAC49239	Aac49239 Arabidops
35	126.2	11.7	483	10 ADE82083	Ade82083 Arabidops
36	125.2	11.6	1155	10 ART23286	Art23286 Seed deve
37	111.4	10.3	121	12 ADG00027	Adg00027 Nicotiana
38	108.2	10.0	1550	8 ADA26540	Ada26540 Arabidops
39	107.4	10.0	541	10 ABX57306	Abx57306 Arabidops
40	105.4	9.8	1371	8 ADA26544	Ada26544 Arabidops
41	103.2	9.6	1815	3 AAC44497	Aac44497 Zea maye
42	102	9.5	860	8 ADA71305	Ada71305 Rice gene
43	95.6	8.9	1665	8 ADA26542	Ada26542 Arabidops
44	94.8	8.8	783	8 ADA70766	Ada70766 Rice gene
45	92.6	8.6	771	3 AAC43020	Aac43020 Arabidops

ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
PI Klessig DF, Kumar D;  
XX  
DR WPI; 2003-278578/27.  
XX  
P-PSDB; ADA26499.  
XX  
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.  
XX  
PS Claim 1; Fig 4; 99pp; English.  
XX  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the SABP2 protein from tobacco plants.  
XX

SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1079; DB 8; Length 1079;

Best Local Similarity 100.0%; Pred. No. 1e-279;

Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGGGAAAGAAAGAACTAACAGGCATATAAATTCAAATGAAGGAGGAAACACT 60  
DB 1 ACGCGGGGAAAGAAAGAACTAACAGGCATATAAATTCAAATGAAGGAGGAAACACT 60  
QY 61 TTGTTTATTAGTACATGCTGCCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
DB 61 TTGTTTATTAGTACATGCTGCCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
QY 121 TAGAGCTCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTCTTGGCACTGATTTGA 180  
DB 121 TAGAGCTCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTCTTGGCACTGATTTGA 180  
QY 181 GAAATAATAGAGGAGCTTCCACACTTTATGATATATCTTTGCGCATGATGGAGTTGATGG 240  
DB 181 GAAATAATAGAGGAGCTTCCACACTTTATGATATATCTTTGCGCATGATGGAGTTGATGG 240  
QY 241 AATCTCTTTACAGAGATGAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAATT 300  
DB 241 AATCTCTTTACAGAGATGAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAATT 300  
QY 301 TGGGACTTGTCTATGGAAAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
DB 301 TGGGACTTGTCTATGGAAAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
QY 361 TCATGCTGATTCGTGTTCACTCCTCTCTTTGTTTGGAAACAGATATAATGACGCGACGC 420  
DB 361 TCATGCTGATTCGTGTTCACTCCTCTCTTTGTTTGGAAACAGATATAATGACGCGACGC 420  
QY 421 CAGCCGAGAAATGGTTGGATACCTCAGTCTTTTACCATATGTTCCCTGAAGGCCACTGA 480  
DB 421 CAGCCGAGAAATGGTTGGATACCTCAGTCTTTTACCATATGTTCCCTGAAGGCCACTGA 480  
QY 481 CATCCATGTTTTTGGCCCCAAAGTCTTGGCTCACAAGCTCTACCAAGCTATGCTCTCCTG 540  
DB 481 CATCCATGTTTTTGGCCCCAAAGTCTTGGCTCACAAGCTCTACCAAGCTATGCTCTCCTG 540  
QY 541 AGGATCTGCATTAGCATCATATGGTGAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600  
DB 541 AGGATCTGCATTAGCATCATATGGTGAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600  
QY 601 CGAAGCCCAAGTATTTTACAGATGACGCTTTGGATCAGTGAAGAGGTTTACATTGCT 660  
DB 601 CGAAGCCCAAGTATTTTACAGATGACGCTTTGGATCAGTGAAGAGGTTTACATTGCT 660  
QY 661 GCACTGAGGATAAGGCATACCAAGAAAGATTCAGCGATGGCAAAATGACAACTTGGTG 720  
DB 661 GCACTGAGGATAAGGCATACCAAGAAAGATTCAGCGATGGCAAAATGACAACTTGGTG 720  
QY 721 TCATGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCAATGCTATCGGAGCCCAAA 780  
DB 721 TCATGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCAATGCTATCGGAGCCCAAA 780  
QY 781 AACTTTGGCGCTCTCTCTTGGAAATTTGCCATATAATACAACTGATCTCTACATTATGCT 840

DB 781 AACTTTGGCGCTCTCTCTTGGAAATTTGCCATAAATACAACTGATCTCTACATTATGCT 840  
QY 841 TCGTCTCATGTCAAGATTTTCAGTGCAATGCTGTAATTTTCTATTTTTCGACCGCGC 900  
DB 841 TCGTCTCATGTCAAGATTTTCAGTGCAATGCTGTAATTTTCTATTTTTCGACCGCGC 900  
QY 901 ATAACTGCTTTGGCTATTTTAAGGATTCAGTAAATTTTCACTCTCTAGTGTGGAAGCT 960  
DB 901 ATAACTGCTTTGGCTATTTTAAGGATTCAGTAAATTTTCACTCTCTAGTGTGGAAGCT 960  
QY 961 TCCCATAAAGATTTGTTCTGTTTCTCCATTCAGAGTGTGTTATGTTGAGATACCTAAAC 1020  
DB 961 TCCCATAAAGATTTGTTCTGTTTCTCCATTCAGAGTGTGTTATGTTGAGATACCTAAAC 1020  
QY 1021 CGTATCAATCTTGTAATGAACCTTCTTCTTTTGGAAAAAATAAAAAA 1079  
DB 1021 CGTATCAATCTTGTAATGAACCTTCTTCTTTTGGAAAAAATAAAAAA 1079  
RESULT 2  
ADA26530  
ID ADA26530 standard; DNA; 918 BP.  
XX  
AC ADA26530;  
XX AC  
XX 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana AtSB2L1 gene homologous to tobacco SABP2 gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO2003016551-A2.  
XX  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
PA  
XX Klessig DF, Kumar D;  
PI  
XX WPI; 2003-278578/27.  
DR  
XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.  
XX  
PS Claim 65; Page; 98pp; English.  
XX  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding  
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L1 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).  
XX  
SQ Sequence 918 BP; 266 A; 175 C; 221 G; 256 T; 0 U; 0 Other;

Query Match 27.4%; Score 295.2; DB 8; Length 918;  
Best Local Similarity 62.7%; Pred. No. 6.9e-69;  
Matches 494; Conservative 0; Mismatches 288; Indels 6; Gaps 2;  
QY 53 AAAACACTTTGTTTGTAGTACATCGTGCATGCCATGAGGTGGAGTTGGTGTACAAAGCTAAA 112  
DB 64 ACAACATTTTGTACTAGTACATCGTTCGTGCCATGCGCGTGTGTCTGTGTACAAAGTTAA 123  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGTTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
DB 124 GCCGCTGTAGAGCGGTGGGCCACCGCGTAACTGTGTGAGCTTAGCTGCCTCCGGAAT 183  
QY 173 TGA---TTTGAGAAAAATAGAGGAGCTTCGCCACATTTTATGATTATACCTTTGTCATTGAT 229  
DB 184 AGACACACGAGGTGCGTCACTGACATCCCATCGGAACATATCTCGGAGCCATTGAC 243  
QY 230 GGAGTTGATGGAATCTCTTTTACGACATGAGAAAGTTATATTAGTGGGCGATAGTCTTGG 289  
DB 244 GAAGCTCCTGACCTCATTTGCCAAATGATGAAAAGTTGTGCTGTTGGTACAGCTTTGG 303  
QY 290 TGGTATGAATTTGGGACTTGCCTATGGAAGTATCCACAAAGATCTATGCTGCTGTTT 349  
DB 304 TGGCTTGAATTTAGCCATAGCCATGGAAGTTTCCGGAATAATCTGTGCTGTATT 363  
QY 350 TTGGCTGCTTTCATGCTGATTCTGTTCACAACTCTCTCTTGTGTTTGGAAACAGTATAA 409  
DB 364 CTGACTGCTTTCATGCGCGACACCGAACACTCACCATCTCTGCTTGGACAGTTTGG 423  
QY 410 TGAGCGAGCCGACGAGAAATTTGGTGGATCTAGTATTTTACCAATATGGTTCCCGTGA 469  
DB 424 AAGCAACATGCTTCAAGAGCATGGATGGCACCGGAATTCGAACCTTATGGTTCAGACAA 483  
QY 470 AGAGCCACTGACATCCATGTTTTCGCCCAAGTTCTTGGCTCACAACTCTACAGCT 529  
DB 484 TTCCGAGCTGA---GTATGTTTTTGTAGCCCTGACTTCATGAAGTTGGGTCTCTACAGCT 540  
QY 530 ATGCTCTCTCTGAGGATCTTGCATTAGCATCATCTTGGTGAGACCAAGCTCTTTGTTTAT 589  
DB 541 TTCTCCAGTTGAGGATCTTGAACCTGGATTAATTTTAAAGGCCAGGATCGTTATTAT 600  
QY 590 GGAAGACCTTATCGAAGCCCAAGTATTTCACAGATGAACGGTTTGGATCAGTGAAGAGT 649  
DB 601 TAACGATTTATCGAAGTGAATAAATCTTCGGATGAAGATATGGTCTGTTCTCTCGAGT 660  
QY 650 TTACATTTGTGCACTGAGGATGAAGCATACAGAAAGATTTCCACGATGCGCAATTCGA 709  
DB 661 TTTCATAGTGTGTAAAGAGGACAAAGCAATTCAGAAAGAACCGCAGAGATGATGATGA 720  
QY 710 CAACATTTGGTGTCACTGMAAGCAATAGAGATTAAAGGTGCTGATCAGATGGCAATGCTATG 769  
DB 721 TAATTTTCCGGTGAATTTAGTGTAGGATGGAGGAGACAGATCATATGCCAATGTTCTG 780  
QY 770 CGAGCCCAAAAACTTTGGCCCTCTCTCTTTGGAAATTTGCCCATPAATAACAACATGATCTCT 829  
DB 781 CRAAGCTCAGCAACTCAGTGATTACTTCTCTGAAAATTCGGGACAAAATTCGTTTAATCAAA 840  
QY 830 ACATTATG 837  
DB 841 TCTTCATG 848

RESULT 3  
ABZ13515  
ID ABZ13515 standard; DNA; 792 BP.  
XX  
AC ABZ13515;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1320.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.  
XX WO200216655-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US026685.  
XX 24-AUG-2000; 2000US-0227866P.  
XX 26-JAN-2001; 2001US-0264647P.  
XX 22-JUN-2001; 2001US-0300111P.  
XX (SCRI ) SCRIPPS RES INST.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX Identifying a stress condition to which a plant cell has been exposed and  
XX producing plants with increased tolerance to these abiotic stresses.  
XX Claim 144; SEQ ID NO 1320; 577pp + Sequence Listing; English.  
XX The invention relates to identifying a stress condition to which a plant  
XX cell has been exposed, comprising: (a) contacting nucleic acid  
XX representative of expressed polynucleotides in the plant cell with an  
XX array or probes representative of the plant cell genome; and (b)  
XX detecting a profile of expressed polynucleotides in the plant cell  
XX characteristic of a stress response. The method is useful in the  
XX production of transgenic plants, cells and seeds and in producing plants  
XX with increased tolerance to abiotic stress. The present sequence is that  
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
XX in methods of the invention. Note: The sequence data for this patent is  
XX not represented in the printed specification but is based on sequence  
XX information supplied to Derwent by the European Patent Office  
XX  
XX Sequence 792 BP; 217 A; 167 C; 196 G; 212 T; 0 U; 0 Other;  
XX  
XX Query Match 27.1%; Score 292.2; DB 6; Length 792;  
XX Best Local Similarity 63.1%; Pred. No. 4.2e-68;  
XX Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;  
QY 53 AAAACACTTTGTTTGTAGTACATCGTGCATGCCATGAGGTGGAGTTGGTGTACAAAGCTAAA 112  
DB 21 ACAACATTTTGTACTAGTACATCGTTCGTGCCATGCGCGTGTGTCTGTGTACAAAGTTAA 80  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGGTTTACAGCCCTTGTATTTAGCAGCTTCTGGCAC 172  
DB 81 GCCGCTGTAGAGCGGTGGGCCACCGCGTAACTGCTGTGGACTTAGCTGCCCTCGGAAT 140  
QY 173 TGA---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATATCTTTGCCATTGAT 229  
DB 141 AGACACAAAGAGTGCATCACTGACATCCCATGCCAATCACTCGGAGCCATTGAC 200  
QY 230 GGAGTTGATGGAATCTCTTTCAGCAGATGAGAGGTATATTAGTGGGCGCATAGTCTTGG 289  
DB 201 GAAGCTCTGACCTCATTTGCCAATGATGAAGTTGTGCTGTTGTGTACAGCTTTGG 260  
QY 290 TGGTATGAATTTGGGACTTGTCTATGGAAGAGTATCCACAAAAGATCTATGCTGCTGTTTT 349  
DB 261 TGGCTTGAATTTAGCCATAGCCATAGGAAAGTTTCCGAAAAAATCTCTGCTGCTGATT 320  
QY 350 CTGCTGCTTTTCATGCTGATTCGTTCACAACTCTCTCTCTTTGTTTGGAAACAGTATTA 409  
DB 321 CTGACTGCTTTTCATGCGCGACACCGAACTCTCACCATCTCTGCTTTGGACAAGTTGG 380  
QY 410 TGAGCGGACGCCAGCCGAGAAATTTGTTGATCTAGTACTCAGTTTACCATATGTTCCCTGA 469  
DB 381 AAGCAACATGCTTCAAGAGAGCATGGGACCCGAAATTCGAACCTTTATGGTTTCAGACAA 440  
QY 470 AGAGCCACTGACATCCATGTTTTTTTGGGCCAAAGTTTCTTGGCTCACAAAGCTCTACAGCT 529

Db	441	TTCCGGACTGA---GTATGTTTTTTAGCCCTGACTCATGAAGTTGGTCTCTACCACT	497	99US-0134768P.
Qy	530	ATGCTCTCTCGAGGATCTTGCAATTAGCATCATCTTGTTGAGACCAAGCTCTTTGTTTAT	589	99US-0134941P.
Db	498	TTCTCCAGTTGAGGATCTTGAACCTGGATTACTTTTAATGAGCCAGGATCGTTATTTAT	557	99US-0135124P.
Qy	590	GGAAGACCTATCGAAGGCCAAGTATTTTCACAGATGAACGGTTTGGATCAGTGAAGAGAGT	649	99US-0135353P.
Db	558	TAACGATTTATCGAAGATGAAAACTTCTCGATGAAGGATATGGTCTGTTCTCTCGAGT	617	99US-0135629P.
Qy	650	TTACATTTGTCACCTGAGGATAAGGCATACCAGAGAAATCCAGCGATGSCAAATTGA	709	99US-0136021P.
Db	618	TTTCTATAGTGTGTAAGAGCAACAGCAATTCAGAAGAACCCAGAGATGGATGATTGA	677	99US-0136392P.
Qy	710	CAACATTTGCTGCTCACTGAAGCAATAGAGATTAAGGTGCTGATCAGTCGCAATGCTATG	769	99US-0136782P.
Db	678	TAATTTCCGGTGAATTTAGTGTGATGGAGATGGAGAGACAGATCATATGCCAATGTTCTG	737	99US-0137222P.
Qy	770	CGAGCCCCAAAACCTTTGGCGCTCTCTCTTTGGAAATTTGCCCAATAATAC	818	99US-0137502P.
Db	738	CAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAATTCGGCGCAAAATTC	786	99US-0137724P.
RESULT 4				
AAC42344				
ID	AAC42344 standard; DNA; 816 BP.			
AC	AAC42344;			
DT	17-OCT-2000 (first entry)			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.			
KW	Hybridisation assay; Genetic mapping; Gene expression control;			
KW	protein identification; signal transduction pathway; metabolic pathway;			
KW	promoter; termination sequence; ss.			
OS	Arabidopsis thaliana.			
PN	EP1033405-A2.			
PD	06-SEP-2000.			
PF	25-FEB-2000; 2000EP-00301439.			
PR	25-FEB-1999;	99US-0121825P.	18-MAY-1999;	99US-0134768P.
PR	05-MAR-1999;	99US-0123180P.	PR	19-MAY-1999;
PR	09-MAR-1999;	99US-0123548P.	PR	20-MAY-1999;
PR	25-MAR-1999;	99US-0125788P.	PR	21-MAY-1999;
PR	29-MAR-1999;	99US-0126264P.	PR	24-MAY-1999;
PR	01-APR-1999;	99US-0127462P.	PR	25-MAY-1999;
PR	06-APR-1999;	99US-0128234P.	PR	27-MAY-1999;
PR	08-APR-1999;	99US-0128714P.	PR	28-MAY-1999;
PR	16-APR-1999;	99US-0129845P.	PR	01-JUN-1999;
PR	19-APR-1999;	99US-0130077P.	PR	03-JUN-1999;
PR	21-APR-1999;	99US-0130449P.	PR	04-JUN-1999;
PR	23-APR-1999;	99US-0130510P.	PR	07-JUN-1999;
PR	28-APR-1999;	99US-0130891P.	PR	08-JUN-1999;
PR	30-APR-1999;	99US-0131449P.	PR	10-JUN-1999;
PR	30-APR-1999;	99US-0132048P.	PR	10-JUN-1999;
PR	04-MAY-1999;	99US-0132407P.	PR	14-JUN-1999;
PR	05-MAY-1999;	99US-0132485P.	PR	16-JUN-1999;
PR	06-MAY-1999;	99US-0132486P.	PR	16-JUN-1999;
PR	06-MAY-1999;	99US-0132487P.	PR	17-JUN-1999;
PR	07-MAY-1999;	99US-0132863P.	PR	18-JUN-1999;
PR	11-MAY-1999;	99US-0134218P.	PR	18-JUN-1999;
PR	14-MAY-1999;	99US-0134219P.	PR	18-JUN-1999;
PR	14-MAY-1999;	99US-0134221P.	PR	18-JUN-1999;
PR	14-MAY-1999;	99US-0134370P.	PR	18-JUN-1999;









PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	21-OCT-1999;	99US-0160880P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 25.3%; Score 273; DB 3; Length 870;			
Best Local Similarity 60.3%; Pred. No. 6.3e-63;			
Matches 488; Conservative 0; Mismatches 315; Indels 6; Gaps 2;			
QY	13	AAAAGAACTAACAGGCATAAATCAATGAGGAAGGAAACACTTTGTTTAGTAC	72
DB	1	AAAAAGTACGNAAGGAAATATGAGTGAGGAGAGAGGAAGCAACACTTCGGTCTATAC	60
QY	73	ATGGTGCATGCCATGAGGTGGAGTTGGTACAAAGCTAAAGCACTGCTAGAGCTGCAG	132
DB	61	ATGGTGGTGGCCAGCGGATGGTGTGTGTACAAAGTTAAGCTCTTCTCGAGGCTTTGG	120
QY	133	GCCATAAGGTTACAGCCCTTGATTTTAGCAGCTTCTGGCACTGA---TTTGAGAAAAATAG	189
DB	121	GCCATCGTTAAACCGCTTTAGACCTAGCTGCTTCCGGTATAGACACAAACAGTCAATCA	180
QY	190	AGGAGCTTCGCACACTTTATGATTATACTTTGCATTGATGAGTTGATGGAATCTCTTT	249
DB	181	CTGACATTTCTACATGTGAACAAATATCTGAGCCATTGATGAGCTTAATGACTTCATTGC	240
QY	250	CAGCAGATGAGAAGGTTATATAGTGGGCGATAGTCTTGGTGGTATGAAATTTGGGACTTG	309
DB	241	CGAATCATGAGAAGGTTGACTCGTTGGTTCATAGCTTTGGAGGTTTGAGTTAGCCTTAG	300
QY	310	CTATGAAAAGTATCCAAAAAGATCTATGCTGCTGTTTCTTTGGCTGCTTTTCATGCGCTG	369
DB	301	CCATGGATAAGTTTCCCGATAAAATCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
QY	370	ATTCCTGTTCACAACT	429
DB	361	ACACCAAACTCACCATCGTTCTGTCGAGGAAAAGTTTGAAGCAGCATGACACCAAGAAG	420
QY	430	ATTGGTTGGATACTCAGTCTTTTACCATATAGTGTTCCTCGAAGCCACTGACATCCATGT	489
DB	421	GATGGATGGCTTGAGCTCGAGACATATGGTT--CAGATAATTCGGCTTGTCTGTGT	477
QY	490	TTTTTGGCCCAAAGTTCTTTGGTTCACAAAGCTTACCAAGTATGCTCTCTCTGAGGATCTTG	549
DB	478	TCTTCAGCACCGACTTTCATGAGCACCGTCTCTACCAACTTCTCTCTGTCGAGGATCTTG	537
QY	550	CATTAGCATCATTTGTTGGAGCAAGCTCTTGTGTTATGGAAGCACTTATCGAAGGCCA	609
DB	538	AGCTTGGATTGCTTCTTAAAGAGCGCCTAGTTCAATGTTTATTAATGAATTCGAAGTGG	597
QY	610	AGTATTTCAGATGAACGGTTTGGATCAGTGAGAGAGTTTACATTGTGTCACCTGAGG	669
DB	598	AGNACTTTTCTGAGAAGGGTATGGATCTGTTCTCTCGAGCTTACATTGTGTGCAAGAGG	657
QY	670	ATAAAGGCATACAGAAAGAAATTCACGAGATGGCAAAATTCACAAATTTGGTGTCTCAGAG	729
DB	658	ACAACATTAATCTCGGAAGACCAATCAACGATGGATGATCCATAATATTCGGCGGAATTTAG	717
QY	730	CAATAGAGATTAAGGTGCTGATCACAATGCGATGCTATGCGAGCCCAAAACTTTGCG	789

Db	718	TGATTGATGGAAGAGACGGATCATATGCCAATGTTTTCGCAACCTCAAGTACTAAGTG	777
QY	790	CCTCTCTCTTGGAAATGCCCCATAAATAC	818
DB	778	ACCATCTATTGGCAATCGCTGACAATTTC	806
RESULT 7			
ID	AAC47519	standard; DNA; 864 BP.	
XX	AAC47519;		
XX	18-OCT-2000	(first entry)	
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 54118.	
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	04-JUN-1999;	99US-0137502P.	
PR	07-JUN-1999;	99US-0137724P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	
PR	10-JUN-1999;	99US-0138847P.	
PR	14-JUN-1999;	99US-0139119P.	
PR	16-JUN-1999;	99US-0139452P.	

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154799P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159633P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.3%; Score 272.6; DB 3; Length 864;

Best Local Similarity 60.8%; Pred. No. 8.1e-63;

Matches 481; Conservative 0; Mismatches 304; Indels 6; Gaps 2;

QY 31 ATAAATTCAATGAGGAGGAACACCTTTGTTTGTAGTACATGTCATGCCATGGAG 90

DB 17 ATATGAGTGAGGAGGAGGAACACCTTTGTTTGTAGTACATGTCATGCCATGGAG 76

QY 91 GTTGGAGTTGGTACAAAGCTAGAGCTAAAGCCACTGTAGAGCTGCAGGCCATAAGGTTTACAGCCC 150



QY 543 GATCTTGCATTAGCATCATCAITGGTGAGACCAAGCTCTTTGTTATGGAAGACCTATCG 602  
|||||  
Db 640 GATCTTGAATTCGCGAAATGTTGGTGAGGCGCAACCCATTGATTAAAGAAAGATATGGCA 699  
QY 603 AAGGCCAAGTATTTCAAGATGAACCGTTTGGATCAGTGAAGAGAGTTTACATFTGTCG 662  
|||||  
Db 700 GAGAGAAGAAGCTTTCAGTGAGGAAGGATACGGATCCGTTACACGTATATTATTATTGTCG 759  
QY 663 ACTGAGGATAAAGGCATACAGAGAAATCCAGCGATGCGCAATTCGACAACTTGGTGTG 722  
|||||  
Db 760 GGAAGAGATCTTGTCACCCGAAGATTACACGCGATCGATGATCGCAACTTCCGCCA 819  
QY 723 ACTGAAGCATAGAGATTAAAGTGCTGATCAATGCGCAATGCTATGCGAGCCGCCAAAAA 782  
|||||  
Db 820 AAGAAGTAAAGAGATCAAGACGCGAGATCATATGCCAATGTTCTCCAGGCTCAACAA 879  
QY 783 CTTTGGCGCTCTCTTTGAAATGTCCTTAATAACAATGA 824  
|||||  
Db 880 CTATGTGCTCTCTTTGGAGATTGCAATAAATATGCTTAA 921

RESULT 9

ABZ13746

ID ABZ13746 standard; DNA; 792 BP.

AC ABZ13746;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 25.1%; Score 271; DB 6; Length 792;  
Best Local Similarity 61.4%; Pred. No. 2.1e-62;  
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;  
QY 55 AACACTTTGTTTGTATACATGCTGATGCGCATGGAGGTTGGAGTTGGTACAGCTAAAGC 114  
|||||  
Db 23 AACACTTTGTTGCTAGTACATGCTGCGTCCACGGCGCATGGTGTGTTCAAGAGTTAAGC 82  
QY 115 CACTCTAGAGCTCTCAGGCCATAAGTTTACAGCCCTTGTATTAGCAGCTTCTGGCACATG 174  
|||||  
Db 83 CTCCTTCTCGAGCTTTGGGCCATCTGTATACCGCCTTAGACCTAGCTGCTCCGGTATAG 142  
QY 175 A---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATATCTTTGCCATTTGATGG 231  
|||||  
Db 143 ACACAACAGGTCATCACTGACATTTCTACATGTGAACAATATTCTGAGCCATTGATGC 202  
QY 232 AGTTGATGGAATCTCTTTCAGCAGATGAGAGGTTTATATTAGTGGGGCATAGTCTTGGTG 291  
|||||  
Db 203 AGCTAATGACTTTCATTGCGGAATGATGAGAAGGTTGTACTCGTTGGTCATAGCTTTGGAG 262  
QY 292 GTATCAATTTGGGACTTGTCTATGGAAAGTATCCACAAAAGATCTATGCTGCTGTTTCT 351  
|||||  
Db 263 GTTGAGTTTACCTTTAGCCATGGATAAGTTTCCGATAAATCTCTGCTCTGCTTCTG 322  
QY 352 TGGCTGCTTTCATGCTGATTTCTGTTCAAACTCTCTCTCTTTTGGAAAGTATAATG 411  
|||||  
Db 323 TGACTGCAATTCATGCCGACACCAAACTCACCATCTGTTGCTGAGGAAAGTTTGCAA 382  
QY 412 AGCGAGCGCAGCGAGAAATGGTTGGATATCTCAGTCTTTTACCATATAGTTTCCCTGAAAG 471  
|||||  
Db 383 GCAGCATGACACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATA 439  
QY 472 AGCCACTGACATCCATGTTTTTGGCCCAAGTTCTGCTCCACAAGCTCTACCAAGCTAT 531  
|||||  
Db 440 ATTCCGGCTTGTGCTGTTCTTCAGCACCGACTTCATGAAGCACCCTCTCTACCAACTTT 499  
QY 532 GCTCTCTGAGGATCTTGCATTTAGCATCATCTGTTGAGAGCAAGCTCTTTGTTTATGG 591  
|||||  
Db 500 CTCCTGTGGAGATCTTGAGCTTGGATTGTTCTTAAAGAGGCTTAGTTCAATGTTTATTA 559  
QY 592 AAGACCTATCGAAGCCCAAGTATTTTCAAGATGAACGGTTTGGATCAGTGAAGAGATTT 651  
|||||  
Db 560 ATGAATTTATCGAAGATGGAGAACTTTTCTGAGAAAGGGTATGGATCTGTTCTCGAGCTT 619  
QY 652 ACATTGTGCTGACTGAGCATAAAGCATACAGAGAAATTCAGCGATGGGCAAAATGACA 711  
|||||  
Db 620 ACATTGTGTCAAAGAGGACACATTTATCTCGAAGACCATCAAGATGGATGATCCATA 679  
QY 712 ACATTGGTGTCACTGAAGCAATAGAGATTAAGGTGCTGATCACAATGGCAATGCTATGG 771  
|||||  
Db 680 ATTATCCGCGAATTTAGTATTGAGATGGAAGAGACTGATCATATGCAATGTTTGA 739  
QY 772 AGCCCCAAAACCTTTGGCCCTCTCTTTGGAAATTTGCCCATAAATAC 818  
|||||  
Db 740 AACCTCAACTACTAAGTAGCACCATCTATTGGCAATCTGCTGACAAATTC 786

RESULT 10

AAC43133

ID AAC43133 standard; DNA; 792 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 1551; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

Sequence 792 BP; 210 A; 169 C; 187 G; 226 T; 0 U; 0 Other;

PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154P.
XX		PR	02-JUL-1999;	99US-0142055P.
PD		PR	06-JUL-1999;	99US-0142390P.
XX	06-SEP-2000.	PR	08-JUL-1999;	99US-0142803P.
PF		PR	09-JUL-1999;	99US-0142920P.
XX	25-FEB-2000; 2000EP-00301439.	PR	12-JUL-1999;	99US-0142977P.
PF		PR	13-JUL-1999;	99US-0143542P.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624P.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085P.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	PR	28-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	PR	02-AUG-1999;	99US-0146386P.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	24-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	PR	09-AUG-1999;	99US-0147933P.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;			







Db 323 TGACTGCTATGATGCCAGACCAACACACATCACCATCTCTTCGTATGGATAAGCTAAGAA 382  
QY 412 AGCGGACGCGCAGCGAGAAATGGTTGGATACCTAGTGTATTTTACCATATGGTTCCCTCGAAG 471  
Db 383 AAGAACTTCACGAGGAATGGTTAGACACCGTGT---ACGAGGAAACCTTGATT 439  
QY 472 AGCAGTACATCCATGTTTTTTGGCCAAAGTTCTTGGCTCAAAAGCTCTACCGCTAT 531  
Db 440 TTCTAGCGAGTTTGGATTTTGGACCAAGATTCATGGCCAAAGACTTGATCAGTTGT 499  
QY 532 GCTCTCTGAGGATCTTGCAATAGCATCATCATCTGTTGGAGACCAAGCTCTTTGTTATGG 591  
Db 500 CTCAGTCCAAAGATCTGAATGGCGAAATGTTGGAGGGCAAAACCCCATGATTAAAGA 559  
QY 592 AAGACCTATCCAAAGGCCAAGTATTTACAGATGAACGCTTTTGGATCAGTGAAGAGAGTTT 651  
Db 560 AAGATATGGCAGAGAGAGAGAGCTTCAGTGAGAGAGATACGGATCGTTACAGGTATAT 619  
QY 652 ACATTTGTGTCAGTGAAGGATTAAGGCATACCAAGAAATTCACGCGATGGCAAAATTGACA 711  
Db 620 TTATTTGATGCGGAAAGGATCTTGTGTACCCGGAAGATTACAGCGATCGATGATCAGCA 679  
QY 712 ACATTTGTGTCAGTGAAGCATAAGATTAAGGTGCTGATCAGATGGCGAATGCTATGG 771  
Db 680 ACTTTCCCCCAAAAGAGTAATGGAGATCAAGACGCGATCATATGCCAATGTTCTCCA 739  
QY 772 AGCCCCAAAAACTTTGCGCTCTCTCTTGGAAATGCCATAAATACAACTGA 824  
Db 740 AGCCTCAACAACTATGTGCTCTCTCTTGGAGATTGCAAAATAAATATGCGCTAA 792

RESULT 13

ADA26533  
ID ADA26533 standard; DNA; 1109 BP.

XX ADA26533;  
XX

DT 20-NOV-2003 (first entry)  
XX

DE Arabidopsis thaliana AtSB2L4 gene homologous to tobacco SABP2 gene.

XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX W02003016551-A2.

PN 27-FEB-2003.

XX 16-AUG-2002; 2002WO-US026312.

XX 16-AUG-2001; 2001US-0312863P.

PR (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX Klessig DF, Kumar D;  
XX WPI; 2003-278578/27.

DR

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for

PT producing transgenic plants with increased resistance to disease, or for

PT screening SABP2 modulators that confer enhanced resistance of plants to

PT disease.

XX Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid

CC protein, sequences that specifically hybridize with the nucleic acid, the

CC complement of the nucleic acid or a natural allelic variant of the

CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for

CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also

CC useful for producing transgenic plants with increased resistance to

CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful

CC for studying the molecular mechanisms responsible for the modulation of

CC salicylic acid-mediated disease resistance in plants. This sequence

CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L4 which

CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.

CC (Note: this sequence is not given in the specification but is derived

CC from the Genbank accession number given by the inventors).

XX

SQ Sequence 1109 BP; 365 A; 193 C; 235 G; 316 T; 0 U; 0 Other;

Query Match 23.7%; Score 256; DB 8; Length 1109;

Best Local Similarity 58.6%; Pred. No. 2.6e-58;

Matches 464; Conservative 0; Mismatches 325; Indels 3; Gaps 1;

QY 36 ATTCAAATGAAGGAAGGAAAAACATTTGTTTGTAGTACATGGTGCATGCCATGGAGTTGG 95

Db 1 ATGGAGAGAGAACAAACAAGAACGGTTTGTCTCTCGTCCATGGGCTGTGCCACGCGCTTGG 60

QY 96 AGTTGGTCAAGCTAAAGCCACTGTAGAGCTGCAGGCCATAAGTTACACCCCTTGAT 155

Db 61 TGTGTGTACAAAGGTGAAAAACGCATCTGGAGGCTGTAGTCTACTGTGTGACCGCGTGGAT 120

QY 156 TTAGCAGCTTCTGGCAGCTGATTGTGAGAAAAATAGAGGAGCTTCGCACACTTTATGATTAT 215

Db 121 CTAGCTGCATCCGCTATAAATATGACTAGATTGGAGAGATTTCAGACTTTGAAGGATTAC 180

QY 216 ACTTTGCCATTGATGGAGTTGATGGAATCTCTT---TCAGCAGATGAGAGGTTATATTA 272

Db 181 TGCAAACCTTTGCTTTGAGTTACTGAACTCACCTGGCTGGATGACGATAAGGTGATCTT 240

QY 273 GTGGGGCATAGTCTTGGTGGTATGAATTTGGAGCTTGTATGGAAGAGTATCCACAAAG 332

Db 241 GTTGGGCATAGTATGGGAGGAATACCTGCTGCTCTCGTCTTCGACATATTCCTTAGTAAG 300

QY 333 ATCTATGCTGCTGTTTCTTGGCTGCTTTCATGCGCTGATTCTGTTCCACAACCTCTCTTT 392

Db 301 ATGCTACTATTTGTTTCTTGACAGCTTTTATGCCCGCACAAAGAACTACTCTGCTTAT 360

QY 393 GTTTGGAAACAGTATAATGAGCGGACGCCAGCGAGAAATGGTTGGATACTCAGTTTTTA 452

Db 361 GTTTACCAAAAGCTAATCAGAGCGTTCACAAGAAGGATGTTGGACACCGGTTTGA 420

QY 453 CCATATGTTCCCTGAAGAGCCACTGACATCCATGTTTGTGGCCCAAGTTCTTGGCT 512

Db 421 ACGTATGGGAAACATGAATGTCTCTAGAGTTTGTCTTTTGGACCACAAAGTTTCATGGCC 480

QY 513 CACAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCAATAGCATCATCTTGGTGAGA 572

Db 481 AAGAATTTGTATCAACTCTCTCCGGTCCAAGATCTTGAATTTGCCGAAATATGTTGGTGA 540

QY 573 CCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAAGTATTTTCACAGATGAACGGTTT 632

Db 541 GTAAACCCCATCATTTACAAATAATCTGGCAGGACAAGAGCTTTAGTGAGGAAGGTAC 600

QY 633 GGATCAGTGAAGAGAGTTTACATTTGTGCTACTGAGGATAAAGGCATATCCAGAAATTC 692

Db 601 GGTACCGTTACACGTATATATATTTGTATGTGGAGAGGACATGCGCGGTACCCGAGGATTAC 660

QY 693 CAGCGATGCGAAATTTGACAACATTTGCTCACTGAAGCAATAGATTTAAAGGCTCTGAT 752

Db 661 CAGTGTGGATGATCAAGAAGCTTTCCGCCAAAAGAAAGTAAATGGAGATCAAAATGTCAGAT 720

QY 753 CACATGGCAATGCTATGCGAGCCCCCAAAACCTTTGCGCTCTCTCTTGGAAATGCCCCAT 812

Db 721 CATATGGCAATGTTCTCAAGCCTCACAAACTATGTGCTCTCTCTGCGGAGATTGCATGT 780

QY 813 AAATACAACTGA 824

Db 781 AAATATGCTAA 792

RESULT 14  
 ADA26535  
 ID ADA26535 standard; DNA; 1046 BP.  
 XX AC  
 ADA26535;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Arabidopsis thaliana AtSB2L6 gene homologous to tobacco SABP2 gene.  
 XX KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
 XX KW modulator; tobacco; ds; gene.  
 XX XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 OS WO2003016551-A2.  
 PN PN  
 XX PD 27-FEB-2003.  
 XX PF 16-AUG-2002; 2002WO-US026312.  
 XX PR 16-AUG-2001; 2001US-0312863P.  
 XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 XX PI Klessig DF, Kumar D;  
 XX DR WPI; 2003-278578/27.  
 XX XX  
 PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.  
 XX Claim 65; Page; 98pp; English.  
 XX The invention relates to a novel isolated 1079 bp salicylic acid-binding  
 CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening  
 CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L6 which  
 CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
 CC (Note: this sequence is not given in the specification but is derived  
 CC from the Genbank accession number given by the inventors).  
 XX Sequence 1046 BP; 330 A; 195 C; 248 G; 273 T; 0 U; 0 Other;  
 SQ

Query Match 23.4%; Score 252.6; DB 8; Length 1046;  
 Best Local Similarity 58.2%; Pred. No. 2.1e-57;  
 Matches 463; Conservative 0; Mismatches 329; Indels 3; Gaps 1;  
 QY 26 AAGGCATTAATCAATGAAGGAGGAAACACATTTGTTTGTACATGGTGCATGCCA 85  
 DB 12 AAGAAAGGCAATGGAGAAATGAACCAAGGCGGTTTGTGCTCATCGGGGTGGCCA 71  
 QY 86 TGGAGGTTGGAGTTGGTCAAGCTAAAGCCACTGTAGAAAGCTGAGGCCATGAAGTTAC 145  
 DB 72 CGGGGATGAGCTTGGGACAGAGTGAAAGCGAGCTGGAGTTGCAAGTCACTGTGTGAC 131  
 QY 146 AGCCCTGATTTAGCAGCTTCTGGCACTCATTTTGAGAAAAATAGAGAGCTTGCACACT 205  
 DB 132 GGCAGTGGATCTTGTGTCATCAGGTATAACATGACCAAGTGAAGAGATTGAGACTCT 191  
 QY 206 TTATGATTACTTCTGCCATTGATGGAGTTGATGGATCTCTT---TCAGCAGATGAGAA 262

Db 192 GAACGATTACTGCAAAACCATTTGCTTGAGTTTCTGAGCTCGCTTGGCTCAGATGACGGTAA 251  
 QY 263 GGTATATATTAGTGGGCATAGTCTTGGTGGTATGAATTTGGAGCTTCTGCTATGGAAGTA 322  
 Db 252 GGTGATTTGTTGTTGCTCATAGCATGGGAGGAATATCCGCTGCATCTTGTCTGTCAGAGCTT 311  
 QY 323 TCCACAAAAGATCTATGCTGCTGCTTTTCTTGGCTGCTTTTATGCTGCTGATTTCTGTTCAAA 382  
 Db 312 CGCTTGAAGATTCGCGCTATTGCTTTTGGACAGCTTTTATGCGCCGACACAAATAACCC 371  
 QY 383 CTCTCTCTTTGTTTTGGAACAGTATAATGACGGCAGCGCCAGAGAAATGGTTGGATAC 442  
 Db 372 ACCTGCTTATGTTTACGAAAGAGCTGCTCAGAAGCATTCACAAGAGGAATGGTTGGACAC 431  
 QY 443 TCAGTTTTTACCATATGTTTCCCTGAAGAGCCACTGACATCCATGTTTTTTGGCCCAA 502  
 Db 432 CACGTGTGTGAACCTACGGGAAACCTGATTTTCTCTACAGTATACTCTTTTGGGACCAA 491  
 QY 503 GTTCTTGGCTCACAGCTCTACAGCTATGCTCTCCTGAGGATCTTGCAATAGCATATC 562  
 Db 492 GTTTATGGCCAAAGAAATGTTATCAAACTCTCCAGTTCAAGATCTTGAAGTGGTGAAGAC 551  
 QY 563 ATTGCTGAGACCAAGCTCTTTTGTATGGAAGACCTATCGAAGGCCAAGTATTTTCACAGA 622  
 Db 552 ATTAGTAGGGGAAACCGTTAGTTACAAACAATCTGGCAGGGACAGAAGCTTTAGTGA 611  
 QY 623 TGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGTGCTGCTGAGGATAAGGCATACC 682  
 Db 612 GGAAGGGTACCGATCCGTTTACACGTATATATTTGTATGTCAGAGAGGATCTTGTGGAAGT 671  
 QY 683 AGAAGAAATTCAGCGATGGCAAAATTTGACAACTTGGTGTCTACTGAAGCAATAGAGATTA 742  
 Db 672 CGAAGATTACCAGCGTTGGATGATTAGCACTTTCCACCAAAAGAGTAATGGAGATCAA 731  
 QY 743 AGGTGCTGATCACATGGCAATGCTATGCGAGCCCAAAACCTTTTCCGCTCTCTCTTTGGA 802  
 Db 732 ATGTGCGATCATATGCCAATGTTCTCCAAGCCCTCAAGAAGTTTGTGCTCTTCTCTGGA 791  
 QY 803 AATTGCCCAATAA 817  
 Db 792 GATTGCAATAAATA 806  
 RESULT 15  
 ADA26538  
 ID ADA26538 standard; DNA; 928 BP.  
 XX AC  
 ADA26538;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Arabidopsis thaliana AtSB2L9 gene homologous to tobacco SABP2 gene.  
 XX KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
 XX KW modulator; tobacco; ds; gene.  
 XX OS Arabidopsis thaliana.  
 PN WO2003016551-A2.  
 PD 27-FEB-2003.  
 XX PF 16-AUG-2002; 2002WO-US026312.  
 XX PR 16-AUG-2001; 2001US-0312863P.  
 XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 XX PI Klessig DF, Kumar D;  
 XX DR WPI; 2003-278578/27.  
 XX XX

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.  
XX  
XX  
PS Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L9 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).

XX SQ Sequence 928 BP; 279 A; 174 C; 214 G; 261 T; 0 U; 0 Other;

Query Match 22.5%; Score 243.2; DB 8; Length 928;

Best Local Similarity 57.0%; Pred. No. 6.7e-55;

Matches 465; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY	11	AGAAAGAACTAACAGGCGATAAATTCAAATGAAGGAAGGAAAAACATTTTGTATTAGT	70
DB	11	AAAAAAGGCAAACTACAGAAAGAAAGAAACGGAGAGATGAAGCATTTATGTCTAGT	70
QY	71	ACATGTGCATGCCATGGAGTTGGAGTTGGTACAGCTTAAGCCACTGCTAGAGCTGC	130
DB	71	TCACGGAGGCTGCCCGTGCCTGTTGGTACAAAGGTGAAGCCGCTTTGAACATTC	130
QY	131	AGCCCATAGGTTACAGCCCTTCATTTAGCAGCTTCTGGCAGCTGATTCGAGAAAAATAGA	190
DB	131	CGGCCACCGTGTACCGGTTTTTGATCTTAGCGCGCATGGTGTGAACATGAGCAGATAGA	190
QY	191	GGAGCTTCGCACACTTTATGATTAATCTTTGCCATTTGATGGAGTTGATGGAATCTCTTTC	250
DB	191	AGATATTCAGACTTTTGAGAGATTTGCTAAGCCGTTGCTTGAGTTCTTGAGTCTTTTGG	250
QY	251	AGCAGATGAGAGGTTATATTAGTGGGGCATAGTCTTTGGTGGTATGAAATTTGGGACTTGC	310
DB	251	CTCGATGATAAAGTAGTCTCTGTCGCGCATAGCCTCGGTGGAATACCGGCTGCTCTTGC	310
QY	311	TATGGAAGATATCCACAAAGATCTATGCTGCTGTTTTCTTGGCTGTTTCATGCTGA	370
DB	311	AGCGGACATGTTTCTTAGTAAATCTCTGTGCTGTTCTCGTTACTTCTTTATGCCGA	370
QY	371	TTCTGTTCAACACTCCTCTTTGTTTGGAAACAGTATAATGAGCGGACCGCCGAGAA	430
DB	371	CACAACGAATCCACCTTCTTACGTGTTGAAAGTTTCTCGAAGCATTCACAGAGAAGA	430
QY	431	TTGGTTGGATATCTAGTTTTTACCATATAGGTTCCCTGAAAGAGCCACTGACATCCATGTT	490
DB	431	ACGTATGGACTTCGAGTTAGGGAGCTATGGAACAGATGACCATCCACTAAAGACTGCTTT	490
QY	491	TTTTGGCCCAAGTCTCTGGCTCACAAAGCTCTACAGCTATGCTCTCTGAGGATCTTGC	550
DB	491	TCTTGGACCTACTACTTG---AAGAAATGATATCTACTTTCTCTATCGAAGATTATGA	547
QY	551	ATTAGCATCATCTGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTTATCGAAGGCCAA	610
DB	548	ATTGGCCAAAATGTTGATGAGAGTCACACGGCTATTACTAGTAATCTGACGGGACTAA	607
QY	611	GTAATTCACAGATGAACGGTTTGGATCAGTGAAGAGATTTTACATTTGTGCTGACTGAGGA	670
DB	608	AAGCTTAACGGCACAGGATATGGATCGATTAGTCGTGTATATCGTATCGGAGAGA	667

QY	671	TAAAGGCATACCAAGAAGATTCCAGCGATGGCAAAATTGACAAACATTTGGTGTCACTGAAGC	730
DB	668	TAAAGGTATACGTGTAGATTTTCCAACGATGGATGATTGAGAACTCTCCGGTTAAGAAGT	727
QY	731	AATAGAGATTAAAGGTGCTGATCATATGGCAATGCTATCGAGCCCAAAAACTTTGCGC	790
DB	728	GATGGAGATCAAAAGATGCAGATCATATGCTATGTTTTCCAAGCCTCATGAACCTCTGTGA	787
QY	791	CTCTCTCTTGGAAATTTGCCATAAATACAACTGATC	826
DB	788	TCGTCTTCTTAAAGATTGCTGATAAATATCCCTAAGC	823

Search completed: September 12, 2005, 02:21:38  
Job time : 597.255 secs

***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 00:50:21 ; Search time 3440.23 Seconds  
(without alignments)  
11938.557 Million cell updates/sec

Title: US-10-780-002-1

Perfect score: 1079

Sequence: 1 acgcggggaagaagaagaa.....gaaaaa 1079

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsa1:\*

9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	575.6	53.3	922	7	CK269872	EST715950
2	566.6	52.5	886	7	CK260266	EST706344
3	562	52.1	914	7	CK270870	EST716948
4	562	52.1	966	7	CK260265	EST706343
5	557.8	51.7	907	7	CK266902	EST712980
6	529.8	49.1	749	5	BQ514677	EST622092
7	523.4	48.5	794	4	BM407979	EST582306
8	521	48.3	781	4	BG126704	EST472350
9	515.2	47.7	784	4	BI434774	EST537535
10	512	47.5	742	7	CK266903	EST712981
11	443.6	41.1	639	5	BQ111509	EST597085
12	429	39.8	646	2	AW621893	EST312691
13	414.4	38.4	590	4	BI928709	EST548598
14	405.6	37.6	535	2	BE433359	EST399888
15	376.4	34.9	957	7	CV267710	WS02032.B
16	371.6	34.4	494	2	BB924354	EST428123
17	363.4	33.7	476	1	AI772823	EST253923
18	358.8	33.3	525	1	AI775631	EST256731
19	351.8	32.6	821	6	CB290604	UCRCS01_0
20	351	32.5	790	7	CK260267	EST706345
21	350.4	32.5	865	7	CB190989	UCRCS06_0
22	350.2	32.5	797	6	CB290370	UCRCS01_0
23	348	32.3	804	6	CB293837	UCRCS01_0
24	346.8	32.1	781	7	CN190896	UCRCS06_0

25	346.6	32.1	811	6	CB292675	UCRCS01_0
26	345.2	32.0	850	6	CB292674	UCRCS01_0
27	342.2	31.7	797	7	CF837566	UCRCS03_0
28	342.2	31.7	834	7	CN191164	UCRCS06_0
29	339.4	31.5	591	4	BI208416	EST526456
30	337.8	31.3	751	7	CN191146	UCRCS06_0
31	337	31.2	747	7	CN189384	UCRCS06_0
32	334.4	31.0	835	7	CF830288	UCRCS01_0
33	332.2	30.8	740	7	CN191507	UCRCS06_0
34	330.2	30.6	828	7	CF838704	UCRCS03_0
35	328.4	30.4	592	4	BI204316	EST522356
36	326.8	30.3	490	7	CK270871	EST716949
37	326.8	30.3	592	4	BI204332	EST522372
38	323.6	30.0	718	5	BQ853353	QGB20E12.
39	323.4	30.0	726	7	CF831499	UCRCS01_0
40	322.4	29.9	726	7	CF831872	UCRCS01_0
41	322	29.8	546	2	AW224223	EST301130
42	319.4	29.6	793	7	CF838361	UCRCS03_0
43	317.2	29.4	712	5	BQ852405	QGB17024.
44	316.8	29.4	420	1	AI772824	EST253924
45	315	29.2	796	7	CF808374	psHB034XM

ALIGNMENTS

RESULT 1  
CK269872  
LOCUS CK269872 922 bp mRNA linear EST 03-AUG-2004  
DEFINITION EST715950 potato abiotic stress cDNA library Solanum tuberosum CDNA  
clone POACP55 5' end, mRNA sequence.  
ACCESSION CK269872  
VERSION CK269872.1 GI:39826850  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Buell,C.R., Hart,A., Ziesmann,V., Karamycheva,S.A. and Baker,B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST715951  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
source  
1..922  
Location/Qualifiers  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POACP55"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DHI08-Tona"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d





QY	69	GTACATGGTGCATGCGCATGGAGGTTGGAGTTGGTCAAGCTAAAGCCACTGCTAGAGCT	128
Db	74	GTACATGGTGCATGCGCATGGAGGTTGGTGTGGTATAAGCTAAAGCCATTTGTAGAGCT	133
QY	129	GCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGAGAAAAATA	188
Db	134	GCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGAGAAAAATA	193
QY	189	GAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTTGATGAGTTGATGGAATCTCTT	248
Db	194	GAGGAACTTTGCACACTTTGATTTATATACCGTGCCATTTGATGAGTTTATGGAATGCTT	253
QY	249	TCAGCAGATGAGAGGTTATATTAGTGGGCACTAGTCTTTGGTGGTATGAATTTGGGACTT	308
Db	254	CCACAAGAGAGAGGTTATCTAGTGGGCACTAGTCTTTGGTGGTATGAATTTGGGACTT	313
QY	309	GCTATGGAAAGATCCACAAGATCTATGCTGCTGCTTTCTTGGCTGCTTTCTGCTTCTGCT	368
Db	314	GCTATGGAAAGATCCACAAGATCTATGCTGCTGCTTTCTTGGCTGCTTTCTGCTTCTGCT	373
QY	369	GATTTCTGTTCCAACTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGAGCGCCGAG	428
Db	374	GATTTCTGTTCCAACTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGAGCGCCGAG	433
QY	429	AATTTGTTGGATCTCAGTTTATTTACCATATGTTTCCCTGAAAGAGCCACTGACATCATG	488
Db	434	GATTTGTTGGATCTCAGTTTATTTACCATATGTTTCCCTGAAAGAGCCACTGACATCATG	493
QY	489	TTTTTTGGCCAAAGTTCTTGGCTCACAAGCTCTACAGCTATGCTCTCTGAGGATCTT	548
Db	494	CTTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTTATCAGTTATGCTCTCTGAGGATCTT	553
QY	549	GCATTAGCATCATTTGTTGAGCAACCAAGCTCTTTGTTTATGGAAGCACTATCGAAGGCC	608
Db	554	TCATTAGCATCTGTTGTTGAGCAACCAAGCTCTCTGTTTATGGAAGCACTATCGAAGGCC	613
QY	609	AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTTGCACTGAG	668
Db	614	AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTATATATCTGCACTGAG	673
QY	669	GATAAAGGATATCCAGAGAAATTCAGCGATGCGAAATTTGACAACTTGG---TGTCAT	725
Db	674	GATAAAGGATATCCAGAGAAATTCAGCGATGCGAAATTTGACAACTTGG---TGTCAT	733
QY	726	GAGCAATAGATTAAGAGTGTCTGATCATGCGCAATGCTATGCGAGCCCAAAACCTT	785
Db	734	GAGCAATAGATTAAGAGTGTCTGATCATGCGCAATGCTATGCGAGCCCAAAACCTT	793
QY	786	TGCGCCTCTCTCTTGGAATTTGCCATAAATCAAACTGATCTTACATTTATGTTCTGTC	845
Db	794	TGTCGCACTCTCTTGGAATTTGCCATAAATCAAACTGATCTTACATTTATGTTCTTAC	852
QY	846	TCATGTCAGATTTTCAGTGCATGCTGTAATTT	878
Db	853	TCCTTTGATAAATAGTAGTTGTTGCTTTCACTT	885
RESULT 3			
CK270870			
LOCUS			
DEFINITION			
EST716948 potato abiotic stress cDNA library Solanum tuberosum cDNA			
clone POACV79 5' end, mRNA sequence.			
CK270870			
CK270870.1 GI:39827848			
EST.			
SOURCE			
Solanum tuberosum (potato)			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
asterids; lamids; Solanales; Solanaceae; Solanum.			
REFERENCE			
1 (bases 1 to 914)			
AUTHORS			
Buell, C.R., Hart, A., Ziemann, V., Karamycheva, S.A. and Baker, B.			
TITLE			
Generation of ESTs from abiotic stressed potato tissue			

JOURNAL  
COMMENT

Unpublished (2003)  
Other ESTs: EST716948  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>  
Seq primer: ANT TAG GTG ACA CTA TAG.

FEATURES

source

1..914  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POACV79"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,  
2d). Set 2 were grown under the standard conditions and then  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 3 were grown under the standard conditions and  
then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

ORIGIN

Query Match 52.1%; Score 562; DB 7; Length 914;  
Best Local Similarity 80.2%; Pred. No. 4.3e-134;  
Matches 685; Conservative 0; Mismatches 165; Indels 4; Gaps 2;  
QY 9 AAAGAAAAAGAACTAACAGGCGATATAAATTCAAATGAAGGAAAGGAAAAACACTTTGTTTAA 68  
Db 51 AAAGAAAGAAAAAAACACAGCCAAAAATGGAGGTTATGAGAAACACTTTGTTTAA 110  
QY 69 GTACATGGTGCATGCGCATGGAGTTGGTGTGGTCAAGCTAAAGCCACTGCTAGAGCT 128  
Db 111 GTACATGGTGCATGCGCATGGAGTTGGTGTGGTATAAGCTAAAGCCACTTTAGAGGCT 170  
QY 129 GCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGCGCACTGATTTGAGAAAAATA 188  
Db 171 GCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGCGCACTGATTTGAGAAAAATA 230  
QY 189 GAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTTGATGAGTTGATGGAATCTCTT 248  
Db 231 GAGGAGCTTCGCACACTTTGATTTATACCGTGCCATTTGATGAGTTGATGGAATGCTCTT 290  
QY 249 TCAGCAGATGAGAGGTTATATTAGTGGGCACTAGTCTTTGGTGGTATGAATTTGGGACTT 308  
Db 291 CCACAAGAGGAGAGGTCATCTAGTGGGCGCATAGTTATGTTGGTATGAATTTGGGACTT 350  
QY 309 GCTATGGAAAGATCCACAAGATCTATGCTGCTGCTTTCTTGGCTGCTTTTCAGCTCT 368  
Db 351 GCTATGGAAAGATCCACAAGATCTATGCTGCTGCTTTCTTGGCTGCTTTTCAGCTCT 410  
QY 369 GATTCTGTTTCACTCTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGAGCGCCGAG 428





ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 749)  
AUTHORS Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,  
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J., and  
Karamycheva, S.A.  
TITLE Generation of a set of potato cDNA clones for microarray analyses  
JOURNAL Unpublished (2002)  
COMMENT Other ESTs: EST622091  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: 17.  
FEATURES  
Location/Qualifiers  
1..749  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec or Binjete"  
/db\_xref="taxon:4113"  
/clone="STMIN36"  
/tissue\_type="mixed tissues"  
/lab\_host="SOLR"  
/clone\_lib="Generation of a set of potato cDNA clones for  
microarray analyses mixed potato tissues"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Combination of untreated and Phytophthora  
infestans-treated libraries of stolons, leaves, leaflets,  
axillary buds of stem explants, petioles, germinating  
eyes, tubers, or roots."  
ORIGIN  
Query Match 49.1%; Score 529.8; DB 5; Length 749;  
Best Local Similarity 82.6%; Pred. No. 8.9e-126;  
Matches 619; Conservative 0; Mismatches 127; Indels 3; Gaps 1;  
QY 61 TTGTTTGTAGTACATGGTGCATGCCATGGAGTTGGAGTTGTACAGCTAAAGCCACTGC 120  
DB 1 TTGTTTGTATGATGTGTATGTCATGGAGTTGGTGTATAGCTAAAGCCATTGT 60  
QY 121 TAGAAGCTCAGGCCATAAGGTTACAGCCCTTGATTAGCAGCTCTGGCACTGATTTGA 180  
DB 61 TAGAGCTCGCGCCACACGGTCACTGCCCTTGACATGCCGCCCTGGCAATTGTTGA 120  
QY 181 GAAATAAGAGGAGCTTCGCACACTTTATGATTATACCTTTGTCATTTGATGGATGG 240  
DB 121 GAAATAAGAGGAACTTTGCACACTTTGATTATACCTGGCAATTGATGGATTTATGG 180  
QY 241 AATCTCTTTACAGAGATGAGAGTTATATTAGTGGGCATAGTCTTGGTGGTATCAATT 300  
DB 181 AATGCCCTTCAAGAGGAGAGGTCATAGTGGGCATAGTTATGGTGGTATGAATT 240  
QY 301 TGGGACTTGTATGGAAAAGTATCCAAAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
DB 241 TGGCACTTGTATGGAAAATACCCAAAAAGATCTTGTGCTGTTACTTGTACTGCTA 300  
QY 361 TCATGCTGATTTCTGTTCTCACTCTCTCTTTGTTTGGAAACAGTATATATGACGGACGC 420  
DB 301 TTCTGCTGATTTCTACTCATGTCTCTCTATGTTTGGATTAAGTACTTTGGAGACGACAA 360  
QY 421 CAGCCAGAGATTGGTTGGATCTCAGTTTTTACCATATGTTCCCTGGAAGCCACTGA 480  
DB 361 CAAAAGAGGATTGGCTTGTGACACCCCAATTGTATCATATGTTACCCCTGAAAGCCCTCGA 420  
QY 481 CATCCATGTTTTTTTGGCCCAAGTTCTTGGCTCACAAAGCTCTACCAAGCTATGCTCTCTG 540  
DB 421 CATCCATGCTTTTTTGGCCCCAGTTCTTGGCTCAAAAACCTTTATCAGTTATGCCCTCTG 480

QY 541 AGGATCTTGCAATAGCATCATCTTGTGTGAGCAACAGCTCTTTGTTATGGAAGACCTAT 600  
DB 481 AGGATCTTTTATTAGCATCGTGTGTGTGAGCAACAGCTCTCTGTTTATCGAAGATCTAT 540  
QY 601 CGAAGGCCCAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGT 660  
DB 541 CAAAGGTGAAGTATTTTACGACGACGAGGTTTGGATCAGTGAAGAGAGTTTATATAATCT 600  
QY 661 GCATGAGGATAAAGGCATACCAAGAATTTCCAGCGATGGCAATTTGACAACATTTGG-- 718  
DB 601 GCATGAGGATAAAGATATCACAAAAGAAATTCAGCAATGGCAATTTGACAACATTTGAA 660  
QY 719 -TGTCACTGAAGCAATAGAGATTAAAGTGTCTGATCAGTCGCAATGTCATGCGAGCCCC 777  
DB 661 CCGTCATCGAAGCAAGAGAGATCAAGGTACTGATCATATGGAATGCTATGTATGAGCCCA 720  
QY 778 AAAAACTTTGGCCCTCTCTCTTTGGAATTT 806  
DB 721 AACGACTTTGTCGCACTCTCTTTGGAGATT 749  
RESULT 7  
BM407979 794 bp mRNA linear EST 10-MAR-2003  
LOCUS EST582306 potato roots Solanum tuberosum cDNA clone cPRO32J8 5'  
DEFINITION end, mRNA sequence.  
ACCESSION BM407979  
VERSION BM407979.1 GI:18259609  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 794)  
AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S.,  
Utterback, T., Chiemiango, A., Bougri, O., Buell, C.R., Ronning, C.,  
Tanksley, S., and Baker, B.  
Generation of ESTs from potato roots  
Unpublished (2001)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.  
FEATURES  
Location/Qualifiers  
1..794  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cPRO32J8"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"  
/clone\_lib="potato roots"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Cornell University, Tanksley lab;  
sequencing; The Institute for Genomic Research. Roots were  
isolated from in vitro grown stem cuttings on CM medium.  
Roots were isolated two weeks after placing the stem  
cuttings from in vitro grown plants on medium."  
ORIGIN  
Query Match 48.5%; Score 523.4; DB 4; Length 794;  
Best Local Similarity 80.7%; Pred. No. 4e-124;  
Matches 624; Conservative 0; Mismatches 146; Indels 3; Gaps 1;  
QY 9 AAAGAAAAGAACTAACCAAGGCATAAAAATTCAAATGAAGGAAGGAAAAACACTTTGTTT 68

```
Db 22 AAGAAAGAAAAAACAACAAGCAAAAAATGAGGTATGAAGAAAACTTTGTTTG 81
Qy 69 GTACATGGTGCATGCCATGAGGTTCGAGTTGGTCAAGCTAAAGCCATGCTAGAAGCT 128
Db 82 GTACATGGTGCATGCCATGAGGTTCGAGTTGGTCAAGCTAAAGCCATGTTAGAGCT 141
Qy 129 GCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGAGAAAAATA 188
Db 142 GCGGCCACACGGTCACTGCCCTTGACATGGCGCCTCTGGCATGATTTGAGAAAAATA 201
Qy 189 GAGGAGCTTCGCACACTTATGATATATCTTTGTCCTATGATGAGGAGTTGATGAATCTTT 248
Db 202 GAGGAATTTGGCAGCTTGTGATATACCGTCCATGATGAGGTTTATGGAATGCCCTT 261
Qy 249 TCAGCAGATGAGAAGGTATATTTAGTGGGCGCATAGTCTTTGGTGGTATGAATTTGGGACTT 308
Db 262 CCACAGAGAGAGAGTCACTAGTGGGCGCATAGTTATGGTGGTATGAATTTGGCACTT 321
Qy 309 GCTATGGAAAGTATCCACAAAGATCTATGCTGCTGTTTCTTGCTGCTTTTCATGCT 368
Db 322 GCTATGGAAAGTATCCACAAAGATCTTTGTTGCTGTTTACTTGACTGCTATTTCTGCT 381
Qy 369 GATTCTGTTCAACATCTCTCTTTGTTTGGAAACAGTATAATGAGGGAGCGCCAGCGAG 428
Db 382 GATTCTACTCACATGTCCTCTCTATGTTTGGATAAGTACTTTGGAGACGCAACAAAAAGAG 441
Qy 429 AATTGGTGGTACTACAGTTTATACCATATGTTTCCCTGAGAGCCACTGACATCCATG 488
Db 442 GATTGGCTTGACACCCAAATTTGTATCATATGGTACCCCTGAGAGGCTCGGACATCCATG 501
Qy 489 TTTTGTGGCCAAAGTTCCTGGCTCACAAGCTCTACAGCTATGCTCTCTGAGGATCTT 548
Db 502 CTTTGTGGCCCAAGTTCCTGGCTCACAAGCTCTATCATGTTATGCTCTGAGGATCTT 561
Qy 549 GCATTAGCATCATATGTTGGTGAACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCC 608
Db 562 TCATTAGCATCGTGGTGGTGAACCAAGCTCTCTGTTTATGGAAGATCTATCAAGGTG 621
Qy 609 AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTTGCTACTGAG 668
Db 622 AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTATATATCTGCACTGAG 681
Qy 669 GATAAGGCATACCAAGAAATTCACGCGATGGCAAAATTTGACAAATTTGG---TGTCACT 725
Db 682 GATAAGTATACCAAAAGATTCACGCAATGGCAAAATTTGACCAACCATGGGACCGTCATC 741
Qy 726 GAAGCAATGAGATTAAGAGTGTGATCATCATGGCAATGCTATGCGAGCCCCA 778
Db 742 GAAGCAAGGAGATCAAGAGTACTGATCATATGGCAATGCTATGTAAGCCCCAA 794
```

```
RESULT 8
BG126704
LOCUS
DEFINITION
  EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone
  cTOF13K13 5' sequence, mRNA sequence.
ACCESSION
  BG126704
VERSION
  BG126704.1 GI:12626892
KEYWORDS
  EST.
ORGANISM
  Lycopersicon esculentum (tomato)
  Lycopersicon esculentum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
  1 (bases 1 to 781)
  van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
  Hansen,C., Ronning,C. and Tankeley,S.
  Generation of ESTs from tomato shoot/meristem tissue
  Unpublished (2001)
  Contact: CUGI
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
```

```
FEATURES
  source
  Location/Qualifiers
    1..781
    /organism="Lycopersicon esculentum"
    /mol_type="mRNA"
    /cultivar="TA496"
    /db_xref="taxon:4081"
    /clone="cTOF13K13"
    /tissue_type="shoot/meristem"
    /dev_stage="developing shoots from 4-6wks old plants"
    /lab_host="SOUR"
    /clone_lib="tomato shoot/meristem"
    /note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
    XhoI; Small expanding leaves from the growing tip were
    taken from greenhouse plants (4-6wks old TA496). Tissue
    was immediately frozen in liquid nitrogen."
```

## ORIGIN

```
Query Match 48.3%; Score 521; DB 4; Length 781;
Best Local Similarity 82.6%; Pred. No. 1.7e-123;
Matches 609; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

Qy 42 ATGAAGGAGAAACACCTTTGTTTAGTACATGGTGCATGCCATGGAGTTGGATTGG 101
Db 45 AAGAAACAAGGAAGACACTTTGTTTGGTACATGGTGCATGCCATGGAGTTGGTTGG 104
Qy 102 TACAGCTAAAGCCACTGCTAGAACTGCAGGCCATAGGTTACAGCCCTTGATTAGCA 161
Db 105 TACAAACTAAAGCCCCCTCTAGAGGTTGCAGGCCACAAGGTCACAACCTCTTGATTGGCA 164
Qy 162 GCTTCTGCACTGATTTTGAGAAAAATAGAGAGCTTCCACACTTTATGATTATATCTTTG 221
Db 165 GCCTCTGATTTGATTTGAGAAAAATAGAGCAACTTCACACATTTTCATGATTACACTTTG 224
Qy 222 CCATTGATGAGAGTGTGATGGAATCTCTTTCAGCAGATGAGAAAGTTATATTTAGTGGGGCAT 281
Db 225 CCATTGATGGAATTTGATGGAATCTCTTCCAAAGAGGAGAAAGTCATCTAGTTGGACAT 284
Qy 282 AGTCTGGTGTGATGAATTTTGGAGCTTCTATGGAAGAGTATCCACAAAAGATCTATGCT 341
Db 285 AGTCTGGTGTGATGAATTTGGGACTTGTATGGAAGAAATACCCACAAAAGATCTATGTT 344
Qy 342 GCTGTTTTCTTTGGCTGCTTTCATGCTGATTTCTGTTCACAACTCTCTCTTTTGGAA 401
Db 345 GCTGTTTTCTTTGGCTGCTTTCATGCTGATTTCTATTCATCTGCTCTCTATGCTTGGAT 404
Qy 402 CAGTATATGAGCGGACGCCAGCGAGAAATTTGGTGGTATCTCAGTTTATACCATAATGCT 461
Db 405 CAGTATTTTGAGAGGATGCAAAACAATGAATTTGGTTGGACACCCCAATTTGTATCATATGCT 464
Qy 462 TCCCTGGAAGAGCCACTGACATCCATGTTTTCGCCCAAGTTCTTGGCTCACAAGCTC 521
Db 465 TCCCATGAAGAGCTCTGCGCATTCATATTTTGGGCCCAAGTCTTGGCTTACAACTCTT 524
Qy 522 TACCAGCTATCTCTCTGAGGATCTTGCATTTAGCATCATATTTGGTGAGACCAAGCTCT 581
Db 525 TATCAATTTATGCTCTCTGAGGATTTGCAATTTGATATCATCATTTGGGAGAGCAAGTTCT 584
Qy 582 TTGTTTATGGAAGACCTATCGAAGCCCAAGTATTTTCACAGATGAACGGTTTGGATCAGTG 641
Db 585 TTATTTCTAGAAGATTTGTCAAAATCAAAATTTTGCACAGATGAAGGATATGATGATCAGTG 644
Qy 642 AAGAGAGTTTACATTTGTCAGCTGAGGATAAAGGCATACCAGAGAAATTCACGCGATGG 701
Db 645 AAGAAAGTTTATATAGTATGCACAGATGATAAATCTCTTACCAAAAGAGTTTCAAAAATGG 704
Qy 702 CAATTTGACAAACAT---TGGTGTCACTGAAGCAATAGAGATTAAGAGTGTCTGATCACATG 758
Db 705 CAATTTGACAAATTAATAGCATCATAGAAACAAAGAAATTTGAAGTGTCTGATCATATG 764
Qy 759 GCAATGCTATCGGAGCC 775
Db 765 GCAATGCTAAGTATGCC 781
```

```
RESULT 9
BI434774
LOCUS      784 bp      mRNA      linear      EST 30-APR-2003
DEFINITION  EST537535 P. infestans-challenged potato leaf, compatible reaction
            Solanum tuberosum cDNA clone PFCBU48 5' sequence, mRNA sequence.
ACCESSION  BI434774
VERSION    BI434774.1 GI:15259464
KEYWORDS   EST.
SOURCE     Solanum tuberosum (potato)
ORGANISM   Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE  1 (bases 1 to 784)
AUTHORS   Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.
TITLE     Generation of ESTs from Potato Leaves Challenged with Phytophthora
            infestans, Compatible Interaction
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: M13P-R.
FEATURES   Location/Qualifiers
            source
            1..784
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="PFCBU48"
            /tissue_type="leaf"
            /dev_stage="6 week old"
            /lab_host="SOUR"
            /clone_lib="P. infestans-challenged potato leaf,
            compatible reaction"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; supplier: Cornell University, Fry lab; sequencing:
            The Institute for Genomic Research; Whole plants were
            challenged with 20,000 sporangia/ml of the compatible P.
            infestans isolate US 940480. Leaf tissue was collected at
            3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen
            in liquid nitrogen immediately upon removal. Kennebec
            plants showed first symptoms of infection at 48 hours
            after inoculation. NOTE: We cannot exclude the
            possibility that this sequence is actually derived from
            Phytophthora rather than potato."
ORIGIN
Query Match      47.7%; Score 515.2; DB 4; Length 784;
Best Local Similarity 81.9%; Pred. No. 5.3e-122;
Matches 619; Conservative 0; Mismatches 133; Indels 4; Gaps 2;

QY      54  A A A C A C T T T G T T T A G T A C A T G G T G C A T G G A G G T T G G A G T T G G T A C A A G C T A A A G 113
DB      30  A A A C A C T T T G T T G T A C A T G T G C A T G C C A T G A A G T T G G T T G G T A T A A G C T A A A G 89
QY      114 C A C T G C T A G A A G C T G C A G G C C A T A A G G T T A C A G C C C T T G A T T A G C A G C T T C T G G C A C T 173
DB      90  C C A T G T T A G A G C T G C G G C C A C A C G G T C A C T G C C C T T G A C A T G G C C G C C T C T G G C A T T 149
QY      174 G A T T T G A G A A A A T A G A G A G C T T C G C A C A C T T T A T G A T T A C T T T G C C A T T A T G G A G 233
DB      150 G A T T T G A G A A A A T A G A G A A C T T T G C A C A C T T G T T G A T T A T A C C G T G C C A T T A T G A G 209
QY      234 T T G A T G G A A T C T C T T T C A G C A G A T G A G A A G G T T A T A T T A G T G G G G C A T A G T C T T G G T G G T 293
DB      210 T T T A T G G A A T G C T T C C A C A G A G A A G G T C A T A C T A G T A G T G G G G C A T A G T A T G T G T G T 269

294  ATGAATTTGGACCTTGCTATGGAAGATATCCACAAAAGATCTATGCTGCTGTTTCTTTG 353
    |||||
270  ATGAATTTGGACCTTGCTATGGAAGATATCCACAAAAGATCTTGTGTGTTTACTTTG 329
    |||||
354  GCTCCTTTTCATGCTGCTGATTTCTGTTCACAACCTCCCTCCCTTTGTTTGGAAACAGTATAAATGAG 413
    |||||
330  ACTGCTATTTCTGCTCTGATTTCTACTCACATGCTCCTCTATGTTTGGTAAGTACTTGGAG 389
    |||||
414  CGGACGCCAGCCGAGAAATGGTTGGATACCTCAGTTTTTACCATATGTTCCCTGGAAGAG 473
    |||||
390  ACGACAAACAAAGAGGATTGGCTTTGACACCAATTTGTATCATATGTTACCCCTGAAGAG 449
    |||||
474  CCATGACATCCATGTTTTTTTGGCCCAAGTTCTTGCTCACAAGCTTACCACCTATGTC 533
    |||||
450  CCTCGSACATCCATGCTTTTTTGGGCCCAAGTTCTTGCTCAAAAACCTTTATCACTTATGC 509
    |||||
534  TCTCCTGAGGATCTTGCAATAGCATCATCTGGTGAGACCAAGCTCTTTGTTTATGGAA 593
    |||||
510  CCTCTGAGGATCTTTTATTAGCATCGTCTGGTGAGACCAACCTCTCTGTTTATCGAA 569
    |||||
594  GACCTATCGAAGCCAAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTAC 653
    |||||
570  GATCTATCAAAGGTGAAGTATTTTCAAGGACGAAGGTTTGGATCAGTGAAGAAAGTTTAT 629
    |||||
654  ATTGTGTGCACTGAGGATAAGGATACCAAGAGATTTCCAGCGATGGCAAAATTCACAAC 713
    |||||
630  ATAAATCTGCATCTGAGGATNAAGTATCACCAAAAAGAAATTTCCAGCATGGCAAAATTCACAAC 689
    |||||
714  ATTGG---TGTCACCTGAAGCAATAGAGATTAAGAGTTAAAGGTGCTGATCACATGGCAATCTATGC 770
    |||||
690  AATGGAACCGTCTATCGAAGCAAGGAGATCAAGTACTGATCATATGGAATGCTATGT 749
    |||||
771  GAGCCCCCAAAACCTTTTGGCCTCTCTCTCTTGGAAATT 806
    |||||
750  AAG-CCCAACGACTTTGTGCGACTCTCTTGAGATT 784
    |||||

RESULT 10
CK266903/c
LOCUS      742 bp      mRNA      linear      EST 03-AUG-2004
DEFINITION  EST712981 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POAC751 3' end, mRNA sequence.
ACCESSION  CK266903
VERSION    CK266903.1 GI:39823881
KEYWORDS   EST.
SOURCE     Solanum tuberosum (potato)
ORGANISM   Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE  1 (bases 1 to 742)
AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE     Generation of ESTs from abiotic stressed potato tissue
JOURNAL    Other ESTs: EST712980
COMMENT    Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from the University of Arizona Genomics
            Institute via http://genome.arizona.edu/orders/
            Seq primer: GTA ATA CGA CTC ACT ATA GGG C.
FEATURES   Location/Qualifiers
            source
            1..742
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="POAC751"
            /tissue_type="abiotic stress treated leaf and root tissue"
            /lab_host="DH10B-Tona"
            /clone_lib="potato abiotic stress cDNA library"
```









Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute  
Seq primer: T3.

FEATURES

source  
1. .590  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cTOB26K14"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/notes="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN

Query Match 38.4%; Score 414.4; DB 4; Length 590;  
Best Local Similarity 84.4%; Pred. No. 5.8e-96;  
Matches 466; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
  
QY 42 ATGAAGGAGGAAAAACACTTTGTTTGTAGTACATGTTGCGATGCCATGGAGTTGGAGTTGG 101  
DB 35 AAGAAAACAAGGAAGACACTTTGTTTGGTACATGTTGCGATGCCATGGAGTTGGTGGTGG 94  
  
QY 102 TACAAGCTTAAGCCACTGCTAGAGCTGCAGGCCATAAGSTTACAGCCCTTGATTTAGCA 161  
DB 95 TACAACTAAGCCCTTCTAGAGTTGCAGGCCACAAGGTACAACTCTTGATTTGGCA 154  
  
QY 162 GCTTCTGGCAGCTGATTTGAGAAAAATAGAGGACTTCGCACACTTTATGATTATATCTTTG 221  
DB 155 GCCTCTGGTATTGATTGGAAAAATAGAGCAACTTTCACACATTTTCATGATTACATTTG 214  
  
QY 222 CCATTGATGAGTTGATGAATCTCTTTTACAGAGATGAGAGGTTATATAGTGGGCAAT 281  
DB 215 CCATTGATGGAATTGATGAATCTCTTCCACAAGAGGAGAAAGTTCATCTAGTTGGACAT 274  
  
QY 282 AGTCTTGGTGGTATGAATTTGGGACTTGCTATGGAAGATATCCCAAAGATCTATGCT 341  
DB 275 AGTCTTGGTGGTATGAATTTGGGACTTGTTATGGAAGAAATATCCCAAAGATCTATGTT 334  
  
QY 342 GCTGTTTTCTTGGCTGCTTTTCATGCTGATTTCTGTTTCAAACTCTCTCTTTGTTTGGAA 401  
DB 335 GCTGTTTTCTTGGCTGCTTTTCATGCTGATTTCTATTCATCTCTCTTCTTCTTGGAT 394  
  
QY 402 CAGTATAATGAGCGGACGCCAGCGGAGAAATGGTTGGATATCTAGTTTTCATATGATGT 461  
DB 395 CAGTATTTTGGAGGAGTCAAAACAATGAATGGTTGGACACCAAAATTTGATCATATGT 454  
  
QY 462 TCCCTTGAGAGCCACTGACATCCATGTTTTTTGGCCCAAGTCTTCTGGCTCACAGCTC 521  
DB 455 TCCCATGAGAGGCTCTGCGATCCATATTTTTTTGGGCCCAAGTCTTCTGGCTTACATCTT 514  
  
QY 522 TACCAGCTATGCTCTCTGAGGATCTTGCATTAGCATCATCATTTGGTGGAGACCAAGCTCT 581  
DB 515 TATCAATTATGCCCTCTCAGGATGTTGCAATTGGTATCATCATTTGGGAGACCAAGTCT 574  
  
QY 582 TTGTTTATGGAA 593  
DB 575 TTATTCCTAGAA 586

RESULT 14  
BE433359  
LOCUS  
DEFINITION tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
EST399888 mRNA linear EST 18-MAY-2001

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

clone cLEG13E6, mRNA sequence.  
BE433359  
BE433359.1 GI:9431202  
EST.

Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Lycopersicon esculentum

REFERENCE

AUTHORS

Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romning,C.M.,  
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tanksley,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
Unpublished (2000)

TITLE

JOURNAL

COMMENT

Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES

source

1. .535  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG13E6"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: pBluescriptSMCUadapt; Site 1: EcoRI;  
Site 2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopene accumulation on the blossom end of  
the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

ORIGIN

Query Match 37.6%; Score 405.6; DB 2; Length 535;  
Best Local Similarity 85.2%; Pred. No. 1.1e-93;  
Matches 453; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 34 AAATTCATTAAGCAAGGAAACACTTTGTTTGTAGTACATGCTGCATGCCATGGAGTT 93  
DB 4 AAGCTAAACAAGAAACAAGGAAACATTTCTGTTCTAGTACATGCTGCAGCCATGGAGCTT 63  
  
QY 94 GGAGTTGGTACAAGCTAAAGCCACTGCTAGAGCTGCAGGCCATAAGGTTACAGCCCTTG 153  
DB 64 GGTGTTGGTACAAGCTAAAGCCATTTGCTAGAGCTGCAGGCCACAGGTCCTGCTCTTG 123  
  
QY 154 ATTTAGAGCTTCTGGCAGCTGATTTTGAAAAAATAGAGAGCTTCGCAACACTTTATGATT 213  
DB 124 ACTTAGCAGCTCTGGCATTGATTTTGAAAAAATAGAGCAACTTCACACACTTCATGATT 183  
  
QY 214 ATACTTTGCCATTGATGAGTTGATGGAATCTCTTTACAGCAGATGAGAGGTTATATTAG 273  
DB 184 ACACTTTGCCATTATTTGGAATTAATGGAATCTCTTCCAAAGAGGAGAAAGCCATATTAG 243  
  
QY 274 TGGGCGCATAGTCTTGGTGTATGAATTTGGGACTTGCTATGGAAGATATCCAAAAAGA 333  
DB 244 TTGGACATAGTCTTGGTGGTATGAATTTAGCACTTGCTATGGAAGAAATATCCAAAAAGA 303  
  
QY 334 TCTATGCTGCTGTTTTTCTTGGCTGCTTTTATGCTGCTGATTCTTATTCACATATCTTCTCTATG 393  
DB 304 TCTATGCTGCTGTTTTTCTTGGCTGCTTTTATGCTGCTGATTCTTATTCACATATCTTCTCTATG 363  
  
QY 394 TTTTGGACAGTATATGAGCGGAGCGGAGCCGAGAGATTTGGTGGATCTCAGTTTATAC 453  
DB 364 TTATGGATCAGTACAAATGAACGGACACAGCAGAGAAATTTGGTTAGATCTCAATTTTATAC 423  
  
QY 454 CATATGGTTCCCTGAAGAGCCACTGACATCCATGTTTTTTTGGGCCAAAGTTCTTGGGCTC 513

Db 424 CATATGGTACCCCTGAAGAGCCACATCATCATGACATTTGGTCCCAATTTTGGCTG 483  
QY 514 ACAAGCTTACAGCATGCTCTCTCTGAGGATCTTGCATTAGCATCATCAT 565  
Db 484 ATAAGCTTTACGGTTAAGCCCTCTGGAAGATGTTGCATTACGATTATCAT 535

CV267710 957 bp mRNA linear EST 22-SEP-2004  
WS02032\_B21\_H19 PTXN-IB-N-A-11 Populus balsamifera subsp.  
trichocarpa x Populus nigra cDNA clone WS02032\_H19 3', mRNA  
sequence.

CV267710  
CV267710.1 GI:52520685  
EST.

Populus balsamifera subsp. trichocarpa x Populus nigra  
Populus balsamifera subsp. trichocarpa x Populus nigra  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 957)  
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,  
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,  
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,  
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,  
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and  
Bohlmann, J.

The poplar transcriptome: Analysis of expressed sequence tags from  
multiple cDNA libraries  
Unpublished (2004)  
Contact: Joerg Bohlmann  
Genome BC forest genomics program  
University of British Columbia  
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,  
Vancouver, British Columbia, Canada, V6T 1Z3  
Tel: 1-604-822-0282  
Fax: 1-604-822-6097  
Email: bohlmann@interchange.ubc.ca  
Plate: WS02032 row: H column: 19  
High quality sequence stop: 957.

Location/Qualifiers  
1..957  
/organism="Populus balsamifera subsp. trichocarpa x  
Populus nigra"  
/mol\_type="mRNA"  
/cultivar="NxM6"  
/db\_xref="taxon:293756"  
/clone="WS02032\_H19"  
/sex="Male"  
/lab\_host="E. coli DH10B T1 phage resistant cells"  
/clone\_lib="PTXN-IB-N-A-11"  
/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'  
end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees  
two metres in height and grown under greenhouse conditions  
were exposed to continuous feeding by Cryptorhynchus  
lappathi (poplar and willow borer) adults caged on the  
sapling using mesh bags. Bark (with phloem and cambium  
attached) from within the caged region was harvested 2  
hours, 6 hours and 48 hours after the onset of treatment.  
mRNA was isolated from each tissue source independently  
and equal quantities of mRNA from each tissue were then  
pooled. cDNA was prepared from 5 micrograms of mRNA and  
directionally ligated into the pBluescript II SK (+) XR  
vector using the pBluescript II XR cDNA Library  
Construction Kit according to manufacturer's instructions  
with modifications (Stratagene). Plasmid DNA was then  
transformed by electroporation into DH10B cells  
(Invitrogen) for propagation. Normalization was applied  
according to published methods [Bonaldi M.F. et al. (1996)  
Genome Research 6(9):791] in order to reduce the abundance  
of highly expressed transcripts."

FEATURES  
source

ORIGIN  
Query Match 34.9%; Score 376.4; DB 7; Length 957;  
Best Local Similarity 67.1%; Pred. No. 4.2e-86;  
Matches 533; Conservative 0; Mismatches 261; Indels 0; Gaps 0;  
QY 40 AAATGAAGGAGGAAAAACACTTTGTTTTAGTACATGTCATGCGATGCGATGAGGAGTTTGGAGTT 99  
Db 154 AGACCAAGATCAGNAGCATTTTGTCTAGTACATGAGCTTGCCATGGGGCTTGTGTCT 213  
QY 100 GGTAAGAGCTAAAGCCCACTGCTAGAGCTGAGGCCATAAGGTTACAGCCCTTGATTTAG 159  
Db 214 GGCATAAGTTCAAAAACGCTGCTTGAGTCAGCAAGTAACAGGGTCACGGTGTCTGACCTTG 273  
QY 160 CAGCTTCTGGCACTGATTTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTTACTTT 219  
Db 274 CTGCTTCAGGCGCCAAACATGAAGCAATGCCAAGATGTAGAAACACTTGTATGAATATACGG 333  
QY 220 TGCCATTGATGGAGCTTGTATGGAATCTCTTTCAGCAGATGAGAAGGTTATATTAGTGGGGC 279  
Db 334 AGCTTTTGTAGAGTTTCTGGCCTCATTAACCGAGGAGAGGTCATTTCTAGTAGGGC 393  
QY 280 ATAGTCTTGGTGGTATGAATTTGGGACTTGCTATGGAAGGATATCCACAAAGATCTATG 339  
Db 394 ACAGCTAGGAGGTTTGGATTTGGCTCTTGCTATGGAAGGTTCCCGAGAGAGATTTGCTG 453  
QY 340 CTGCTGTTTTCTGGCTGCTTTTCATGCTGATTTCTGTTTCAAACTCCCTCTTCTGTTTGG 399  
Db 454 TTGCTGTTTTCTTATCAGCTTTTCATGCGAGATACACACAAAGCCATCATTTGTCTTGG 513  
QY 400 AACAGTATAATGAGCGGACCGCAGCGAGAAATTTGGTTGGATATCTCAGTTTTTACCATATG 459  
Db 514 ATCAGTATACGAGAGAGACCCCGGGGATTCCTGGTTGGACACTCAATTTTATACATACA 573  
QY 460 GTTCCCTCGAAGAGCCACTGACATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAGC 519  
Db 574 GCAGTTCTCAAAGTCATCTCACAACAATGTCTTTGGACCCCAAAATTCCTATCTCTCCAAGC 633  
QY 520 TCTACAGCATGCTCTCTGAGGATCTTGCAATTTAGCATCATCATTTGGTGAGACCAAGCT 579  
Db 634 TCTATCAGCTAAGCCCACTGAGGATCTTGAGCAAGCAAAAGACTATGTTAAGGCCAGAT 693  
QY 580 CTTTGTGTTATGGAAGACCTATCGAAGGCCAAGTATTTTCACAGATGAACGGTTTGGATCAG 639  
Db 694 CACTGTTTCTGTATGATTTGTCAAAGGCAACAGTTTCTCCACGCGGCTATGGGTGAG 753  
QY 640 TGAAGAGAGTTTACATTTGTGTGCACTGAGGATAAAGGATACCGAAGAAATTCAGCGAT 699  
Db 754 TCAACAGGATGTATGTTTATCTGCGATGAAGATTTTAGCGATACCGAAGAGATTTCAACGCT 813  
QY 700 GGCAAAATTGCAACATTTGGTCTCACTGAGCAATAGAGATTTAAAGGTCTGATCACATGG 759  
Db 814 GGATGATTGAAACACAGTCTGTTGAAGAAATTTAGAAATTTGAAGGTGCACACCATATGG 873  
QY 760 CAATGCTATGCGAGCCCCCAAAACCTTTTCGCGCTCTCTCTTGGAAATTTGCCCAATAATACA 819  
Db 874 TTAAGTTTCAAGAGCCACCAAGAACTCTTCCATTTGTCTCTCAGAGATACCAATAAACATG 933  
QY 820 ACTGATCTCTACAT 833  
Db 934 CTTAAGCTGTACTT 947

Search completed: September 12, 2005, 06:31:50  
Job time : 3448.23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 01:44:06 ; Search time 205.395 Seconds  
(without alignments)  
8595.849 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaaagaaaa.....gaaaaaaagaaaaaa 1079

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/FCUTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	186	17.2	1078	3	US-08-981-256A-4
2	49	4.5	7218	1	US-08-232-463-14
3	44.8	4.2	1141	4	US-09-806-708B-22
C 4	40	3.7	601	4	US-09-949-016-153422
C 5	40	3.7	601	4	US-09-949-016-153423
6	40	3.7	130563	4	US-09-949-016-12273
7	40	3.7	131379	4	US-09-949-016-16050
C 8	39.2	3.6	83938	4	US-09-949-016-16068
9	38.8	3.6	601	4	US-09-949-016-154474
10	38.8	3.6	1782	4	US-09-248-796A-369
11	38.6	3.6	2475	4	US-09-220-132-179
12	38.6	3.6	2884	3	US-09-439-313-335
13	38.6	3.6	2984	3	US-09-352-616A-335
14	38.6	3.6	2984	3	US-09-232-149A-335
15	38.6	3.6	2984	4	US-09-636-215-335
16	38.6	3.6	2984	4	US-09-685-166A-335
17	38.6	3.6	2984	4	US-09-688-489-335
18	38.6	3.6	2984	4	US-09-679-426-335
19	38.6	3.6	2984	4	US-09-759-143-335
20	38.6	3.6	2984	4	US-09-651-236-335
21	38.6	3.6	60593	4	US-09-949-016-11779
22	38	3.5	32798	4	US-09-949-016-11758
23	38	3.5	32798	4	US-09-949-016-17366
C 24	37.8	3.5	1141	4	US-09-806-708B-22
C 25	37.8	3.5	85850	4	US-09-949-016-13424
26	37.6	3.5	2183	3	US-09-352-990-13
27	37.6	3.5	42000	4	US-10-081-563-25

28	37.6	3.5	392000	4	US-10-027-983-111	Sequence 11, Appl
C 29	37.2	3.4	219964	4	US-09-949-016-15086	Sequence 15086, A
30	37	3.4	832	4	US-09-621-976-2813	Sequence 2813, Ap
31	37	3.4	1929	2	US-08-892-770-2	Sequence 2, Appli
32	36.6	3.4	1793	4	US-09-620-312D-996	Sequence 996, App
C 33	36.6	3.4	7218	1	US-08-232-463-14	Sequence 14, Appl
34	36.6	3.4	23497	4	US-09-949-016-17072	Sequence 17072, A
35	36.6	3.4	149971	4	US-09-949-016-13590	Sequence 13590, A
C 36	36.4	3.4	406	4	US-09-829-481-3	Sequence 3, Appli
C 37	36.4	3.4	1362	4	US-09-710-279-803	Sequence 803, App
C 38	36.4	3.4	1386	3	US-09-134-001C-767	Sequence 767, App
39	36.4	3.4	3553	4	US-09-710-279-4250	Sequence 4250, Ap
40	36.4	3.4	4095	4	US-09-710-279-4079	Sequence 4079, Ap
C 41	36.2	3.4	132	4	US-10-021-338A-66	Sequence 66, Appl
C 42	36.2	3.4	6350	2	US-08-385-335A-9	Sequence 9, Appli
C 43	36	3.3	601	4	US-09-949-016-32018	Sequence 32018, A
C 44	36	3.3	601	4	US-09-949-016-70476	Sequence 70476, A
C 45	36	3.3	60002	4	US-09-949-016-12464	Sequence 12464, A

ALIGNMENTS

RESULT 1  
US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:  
; APPLICANT: Meinhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981-256A  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob  
; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-OZ1112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1078 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-981-256A-4

Query Match 17.2%; Score 186; DB 3; Length 1078;  
Best Local Similarity 53.2%; Pred. No. 3.5e-44;  
Matches 421; Conservative 0; Mismatches 365; Indels 6; Gaps 1;  
QY 38 TCAGTGAAGGAAGAAACACTTTGTTTATGATCGATGCCATGAGGTTTGAG 97  
Db 37 TCAGAAATGGCATTCGCTCATTTTGTCTTATTACCATATGCCACGGTGCATGGAT 96

```
QY 98 TTGGTCAAGCTAAAGCCACTCTCTAGAGCTGCAAGCCATAAGGTTACAGCCCTTGATTT 157
Db 97 TTGGCAAGAGCTCAAAACCCCTCTTTGAGGCACTTGCCCAAGGTTACTGCACTGGACCT 156
QY 158 AGCAGCTTCTGCACTGATTTTGCAGAAATAGAGGAGCTTCCACACATTTATGATTATAC 217
Db 157 TGCAGCAAGCGGCTTGACCCAAAGGCAATGAGGAGATTGGCTCATTTGATGATGATTC 216
QY 218 TTTCGATTTGATGGAGTTGATGGAATCTCTTTTCAGCAGATGAGAAGGTTATATTAGTGGG 277
Db 217 TGAACCTTGTGAGCTTCTTGAGGCACTCCCTCCAGGGGAAAGGTAATTCGTGGTGG 276
QY 278 GCATAGCTTGTGGTATGAATTTGGGACTTCTGATGGAAGGATATCCACAAAAGATCTA 337
Db 277 CGAGAGCTGAGGAGCTCAATATAGCAATTCGCTGATATAATCTGTGAAAGAGATTGC 336
QY 338 TCTGCTGTTTCTTGGCTGCTTTTCATGCTGATTTCTGTTTCAACTCTCTCTTTGTTTTT 397
Db 337 AGCTGCTGTTTTCACAAATTCAGTATTTGCCAGACACCGAGCACTGCCCATCTTACGTCGT 396
QY 398 GGAAACAGTATAATGAGCGGACGCCAGCCGAGAAATGGTTGGATATCTCAGTTTTTATACCATA 457
Db 397 GATTAAGCTCATGGAGGTGTTTCCCG-----ACTGGAAGACACACGATATTTACGTA 450
QY 458 TGGTTCCCTGAAGAGCCACTGACATCCATGTTTTTTTGGCCCAAAAGTTCTTGGCTCAAA 517
Db 451 CACTAAAGATGCAAGGAGATAACTGGATTGAAACTGGGCTTCAGGCTTCTGAGGGAATA 510
QY 518 GCTCTACAGCTATGCTCTCTGAGGATCTTCGATTAGCATCATCATTTGGTGAGACCAAG 577
Db 511 TTTATATACCTTTGCGGCTCTGAGGAATATGAATGGCGAAGATGTTGACAAGGAAGGG 570
QY 578 CTCCTTTGTTTATGGAAGACTATCGAAGGCCAAGTATTTACAGATGAACGGTTTGGATC 637
Db 571 ATCATATTTCAAAATATTTTACGTAAGCGACCATCTTCTCATAGGAAGTTACGATC 630
QY 638 AGTGAAGAGAGTTTACATTTGTGCTGACCTGAGGATAAAGGATACAGAGAAATTCAGCG 697
Db 631 GATTAAGAAATTTATGTGTGACCGACCAAGACGAAATATTTTACCTGAAATTTCAACT 690
QY 698 ATGGCAATTTGACAACTTGTGTCTACTGAAGCAATAGAGATTAAGGTGCTGATCAGAT 757
Db 691 CTGGCAATTAGAAACTATAAACCAAGCAAGGTTTATAAGGTGCGAAGGTGGAGATCAATA 750
QY 758 GGCAATGCTATGCGAGCCCAAAACTTTGCGCTCTCTCTTGGAAATTTGCCCATATAATA 817
Db 751 ATTGCAAGCTTACAAAGACTAAGGAGATCGCTGAATTTCTCCAAGAGGTGGCTGATACCTA 810
QY 818 CAACTGATCTCT 829
Db 811 TAATTGACTTCT 822
```

## RESULT 2

```
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
```

## GENERAL INFORMATION:

```
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
```

## INFORMATION FOR SEQ ID NO: 14:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14
```

Query Match 4.5%; Score 49; DB 1; Length 7218;

Best Local Similarity 7.4%; Pred.No.0.0016; Matches 25; Conservative 177; Mismatches 137; Indels 0; Gaps 0;

```
QY 721 TCACATGAAGCAATAGAGATTAAAGTGCTGATCATGCGCAATGCTATGCGAGCCCAAA 780
Db 1019 TCACAGATTATTTCCGAGCTTGCTCGAGGTGCGAGGAGCTTGGATATTTTTTTTT 1078
QY 781 AACTTTGCGCTCTCTCTTTGGAAATGCGCAATAAACAATGATCTCATATATGCT 840
Db 1079 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1138
QY 841 TCGTCTCATGCAAGATTTTCAGTGCATGCTGTAATTTTTTTCTATTTTCGACGCGCG 900
Db 1139 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1198
QY 901 ATAATGCTTTGCTGCTATTTTAAAGGATTCAGTAAATTTCACTTCTTAGTGAAGCT 960
Db 1199 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1258
QY 961 TCCACATAAGAGATTGCTGTTCTCCATTCCAATGCAAGTGTGTTATGTTGAGATACCTAAAC 1020
Db 1259 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1318
QY 1021 CGTATCAATCTTGTAAATGAACCTTCTTCTTCTCTTTT 1059
Db 1319 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1357
```

## RESULT 3

```
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
```

## GENERAL INFORMATION:

```
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. PAF1 promoters
US-09-806-708B-22
```

```
Query Match          4.2%; Score 44.8; DB 4; Length 1141;
Best Local Similarity 11.8%; Pred. No. 0.01;
Matches 74; Conservative 208; Mismatches 356; Indels 0; Gaps 0;

QY 308 TGTCTATGGAAAGTATCCCAAAAGATCTATGCTGCTGTTTCTTGCTGCTTTCATGCC 367
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 387 KWTBYKRTWNNNGTTTWKEMAWYKMDMDHGTNNNNNGRTYYGHTKMKMT 446
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 368 TGATTCGTTCACAACTCCTCTCTTTGTTTGGAAAGTATAATAGCGGACGCCGCGA 427
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 447 YYKWKANNCKWRANDHKTCTHNNTTWKMKTWNNCYWKSMTNGKSHRBAJAVYTWYMW 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 428 GAATGGTGGTACTCAGTCTTTTACCATATGCTCCCTGAGAGCCACTGACATCCAT 487
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 507 WRYAHANNNDWYWKACTWYBYVCCKWNNYAAWYTKSMNYTSRYRWKTNNSWRM 566
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 488 GTTTTTGGCCCAAGTCTTGCTCACAAGCTCTACAGCTATGCTCTCCTGAGGATCT 547
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 RSDTRSMGRANNYARABHYGKNTRWBSHTWBHBRAGAAYWMBYBAKCHCKWAK 626
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 548 TGCATTAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGC 607
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 627 YRAKYVAGAGGSSNNNNNNNNNNATCARDYVYAAASRWYMANAKWYVYKBAANNA 686
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 608 CAAGTATTTCCAGATGAACGGTTTGCATCAGTGAAGAGAGTTTACATTTGTGCTGCTGA 667
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 687 YTHANNWGCWNNATDTRRTWKNNNNNNNAGTWKNNNNNNNAKNAKNAAYKAAK 746
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 668 GGATAAAGGCATACCAAGAAATTCAGCGATGCGCAATTCACAACTTGGTGCTACTGA 727
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 747 KHRWANKWARGWHDAAABTTDKRNGAYKYITNNNNNNYRGVTVVTAARDGWANN 806
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 728 AGCAATAGAGATTAAGGTGTCATCATCATGCGCAATGCTATGCGAGCCCCCAAAACTTTG 787
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 807 NNNNNNNNNNGSDMWVTVWYVANGYTNNNNNNNNNAYAWWTKWYVTTDDRWBAYT 866
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 788 CGCTCTCTCTGGAAATTCGCCATAAATAACACTGCTCTACATTATGCTCTGCTCTC 847
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 867 NNNNNRMAYGAYADDYAYMSDTCDAWKKWDATKNNNATTYNRGTAWRTNNNNNNMTM 926
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 848 ATGTCAAGATTTTCAGTGCATGCTGTAATTTTCTATTCTATTTTCGACCGCGCATAACTG 907
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 927 KTKYBHAAWNNNNNGKCTAHTWVCATKTKTKGWNCTTTCKYKNNCTWTYWTMT 986
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 908 TCTTTGCCCTATTTAAGGATTCAGTAATTTTCACTCTTCTAGTGTGGAAGGCTTCCACAT 967
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 987 TRTTWYAATRWKTNATGSMTRCNATGKNNNNYWTGK 1024
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 4
US-09-949-016-153422/c
; Sequence 153422, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153422
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-153422

Query Match          3.7%; Score 40; DB 4; Length 601;
Best Local Similarity 48.3%; Pred. No. 0.18;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 848 ATGTCAAGATTTTCAGTGCATGCTGTAATTTTCTATTCTATTTTCGACCGCGCATAACTG 907
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 ATGTTTAATATATATTAATTAAGAACTCTGATGAAGCAGCTGTTTTCGACGAGATATTGGT 175
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 908 TCTTTGCCCTATTTAAGGATTCAGTAATTTTCACTCTTCTAGTGTGGAAGGCTTCCACAT 967
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 ACTTTGAAAATGATAAACATTTCTGAAACTTGCCCTGGCCGAGATACTTATGTTTCTTAAG 115
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 968 AAGGATTTGTTCTGTTTCTCCATTCAAGTGTGTTTATGTTGAGATCTTAAACCGTATCA 1027
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 114 ATAAATTGATTATATCATCTACTGTTTATAAAGTCTTTCTAGAAAACTCAAGTTTGTAA 55
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1028 ATTCCTGTAAAGAACTTCTCTTCTCTCTTTTGAAAAAATAAAAAAAAAAAAA 1079
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 ATAAACTATTGGAATAATTCCTTCTTATGCTTACATTTTAAAAAATACAA 3

RESULT 5
US-09-949-016-153423/c
; Sequence 153423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153423
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-153423

Query Match          3.7%; Score 40; DB 4; Length 601;
Best Local Similarity 48.3%; Pred. No. 0.18;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 848 ATGTCAAGATTTTCAGTGCATGCTGTAATTTTCTATTCTATTTTCGACCGCGCATAACTG 907
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 ATGTTTAATATATATTAATTAAGAACTCTGATGAAGCAGCTGTTTTCGACGAGATATTGGT 289
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 908 TCTTTGCCCTATTTAAGGATTCAGTAATTTTCACTCTTCTAGTGTGGAAGGCTTCCACAT 967
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 ACTTTGAAAATGATAAACATTTCTGAAACTTGCCCTGGCCGAGATACTTATGTTTCTTAAG 229
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 968 AAGGATTTGTTCTGTTTCTCCATTCAAGTGTGTTTATGTTGAGATCTTAAACCGTATCA 1027
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 ATAAATTGATTATATCATCTACTGTTTATAAAGTCTTTCTAGAAAACTCAAGTTTGTAA 169
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



RESULT 9  
US-09-949-016-154474  
; Sequence 154474, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154474  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-154474

Query Match 3.6%; Score 38.8; DB 4; Length 601;  
Best Local Similarity 46.8%; Pred. No. 0.4; Indels 0; Gaps 0;  
Matches 118; Conservative 1; Mismatches 133  
QY 822 TGATCTCTACATATATGCTTCGTCTCATGTCAAGATTTTCAGTGCATGCTGTAATTTTT 881  
DB 148 TTACTTCTACAGTTGTTCTTGGTCACTTACTGCTAATTTATAAGTGTCAAGTCTTG 207  
QY 882 TCTATTTTTCGACCGCGCATAACTGCTCTTGGCTATTTTAAGGATTCGAGTAATTTAC 941  
DB 208 ACCTGCTGTGTGACCACTAATATGTACCTAGCTTTGGCTAAGTCTTTAACTAATTTAT 267  
QY 942 TCTTCTAGTGTGAAGGCTTCCACATAAGGATTTGTTCTGTTCTCCATTCAGTGTGCT 1001  
DB 268 TCTAAAATGAATATAGCTAATATTTACTGAGGCTTGCTATATGCGAGTACTTTCTGA 327  
QY 1002 TATGTTGAGATACTTAAACCGTATCAATTTCTTGTAATGAACCTCTTCTTCTTTTGA 1061  
DB 328 GTGTTTCACATATTTAGACCCATTTATCTCCATGAGTAGGCATTAATTTCTATTTA 387  
QY 1062 AAAAAAAAAA 1073  
DB 388 TAGATAAGAAA 399

RESULT 10  
US-09-248-796A-369  
; Sequence 369, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 369  
; LENGTH: 1782  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-369

Query Match 3.6%; Score 38.8; DB 4; Length 1782;  
Best Local Similarity 64.4%; Pred. No. 0.73; Indels 0; Gaps 0;  
Matches 58; Conservative 0; Mismatches 32  
QY 645 AGAGTTTACATTTGTGTGCACTGAGGATAAGGCATACCAAGAAGATTCAGCGATGGCAA 704  
DB 1219 AAGATATCAATCTTGGACACTGACAATAAAGCACATCACCAGAATTTAAATAGTCAGAAA 1278  
QY 705 ATTGACAACATTTGGTGTCACTGAAGCAATA 734  
DB 1279 ATTAATACAGTGGTGATTAATGATTTCAATA 1308

RESULT 11  
US-09-220-132-179  
; Sequence 179, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANC  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; PRIOR FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 179  
; LENGTH: 2475  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-220-132-179

Query Match 3.6%; Score 38.6; DB 4; Length 2475;  
Best Local Similarity 45.5%; Pred. No. 1; Indels 0; Gaps 0;  
Matches 137; Conservative 0; Mismatches 164  
QY 775 CCCAAAACTTGGCCCTCTCTCTGGAAATGGCCATAAATACAACTGATCTCTACATT 834  
DB 2074 CACTTAAATTTGGAACCTTAAAGTTTAGATGCATTTATATAAAAAACCTTAAAGCAGTATC 2133  
QY 835 ATGCTCTCGTCTCATGTCAAGATTTTCAGTGCATGCTGTAATTTTTTCTATTTCGAC 894  
DB 2134 TGTATTTAGCTGTAAACCAAGTTGGAAGCTATTTCGGATAATTTCTTAAATATGATGAA 2193  
QY 895 CGGCGCATAAAGTCTTTTGGCCTATTTTAAGGATTCAGTAATTTCACTCTTCTAGTGG 954  
DB 2194 CTTTGGAGTACTGTTTCTTCTTCAAACTGAATGTAATTAATTCATGAATAACCT 2253  
QY 955 AAGCTTCCACATAAGGATTTGTTCTTCTTCCATTCAGTGTGTTATGTTGATAC 1014  
DB 2254 TATATGTTTAAACCAATCTTTGTATACATTTTGGGATTTTGGTGTATATGCTAAATCAC 2313  
QY 1015 TTAACCCATCAATCTTGTGAAGAACTCTTCTTCTTCTTTTGAAGAAAAA 1074  
DB 2314 ATTCAGCATGTATTTTGACATTTTAAATACTTCCCTCAATTCGTAAATTTAAAGAAT 2373  
QY 1075 A 1075  
DB 2374 A 2374

RESULT 12  
US-09-439-313-335  
; Sequence 335, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.

```
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-439-313-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTTCTCCATCAAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTTGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTAATGAACCTCTTCTTCTCTTTTGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAAAAATAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 13
US-09-352-616A-335
; Sequence 335, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqi
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTTCTCCATCAAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTTGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTAATGAACCTCTTCTTCTCTTTTGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAAAAATAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 14
US-09-232-149A-335
; Sequence 335, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTTCTCCATCAAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTTGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTAATGAACCTCTTCTTCTCTTTTGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAAAAATAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 15
US-09-636-215-335
; Sequence 335, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
```



```
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-636-215-335

Query Match      3.6%; Score 38.6; DB 4; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCCACATAAGGATTGTTCTGTTCTCCATTCAAGTGTGTTATGTGAGATAC 1014
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2855 AAGTCTGTGACTTGAAGTTTAGTCAGCACCCTCCCAAACTTTATTTTCTATGTGTTT 2914
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1015 TTAAGCGTATCAATTCTTGTAAATGAAACTTCTTCTTTTTCGAAAAA 1074
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2915 TTGCAACATATGAGTGTGTTTGAATAAATAGTACCCTGCTTTATTAGAAAAA 2974
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1075 AAAAA 1079
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2975 AAAAA 2979
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 12, 2005, 06:38:17  
Job time : 208.395 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 04:46:01 ; Search time 2389.95 Seconds  
(without alignments)  
2964.938 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaaagaaagaaaa.....gaaaaaaagaaaaaa 1079

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1079	100.0	1079	21	US-10-780-002-1
2	320	29.7	933	18	US-10-424-599-64748
3	295.2	27.4	989	21	US-10-780-002-32
4	292.2	27.1	792	9	US-09-938-842A-1320
5	292.2	27.1	792	11	US-09-938-842A-1320
6	277.6	25.7	1158	18	US-10-424-599-85023
7	274.6	25.4	870	21	US-10-780-002-33

Sequence 1, Appli  
Sequence 64748, A  
Sequence 32, Appl  
Sequence 1320, Ap  
Sequence 1320, Ap  
Sequence 85023, A  
Sequence 33, Appl

8	271.6	25.2	837	21	US-10-780-002-34	Sequence 34, Appl
9	271	25.1	792	9	US-09-938-842A-1551	Sequence 1551, Ap
10	271	25.1	792	11	US-09-938-842A-1551	Sequence 1551, Ap
11	269	24.9	792	9	US-09-938-842A-217	Sequence 217, App
12	269	24.9	792	11	US-09-938-842A-217	Sequence 217, App
13	268.2	24.9	1035	18	US-10-425-114-12800	Sequence 12800, A
14	256	23.7	792	21	US-10-780-002-35	Sequence 35, Appl
15	252.6	23.4	819	21	US-10-780-002-37	Sequence 37, Appl
16	240.8	22.3	653	18	US-10-424-599-52897	Sequence 52897, A
17	240.8	22.3	834	21	US-10-780-002-39	Sequence 39, Appl
18	239.8	22.2	771	9	US-09-938-842A-1278	Sequence 1278, Ap
19	239.8	22.2	771	11	US-09-938-842A-1278	Sequence 1278, Ap
20	239.8	22.2	771	21	US-10-780-002-40	Sequence 40, Appl
21	232.6	21.6	912	21	US-10-780-002-36	Sequence 36, Appl
22	224.8	20.8	783	21	US-10-780-002-38	Sequence 38, Appl
23	215	19.9	901	21	US-10-780-002-41	Sequence 41, Appl
24	197.8	18.3	920	18	US-10-424-599-128189	Sequence 128189, A
25	195.6	18.1	1101	19	US-10-437-963-62048	Sequence 62048, A
26	191.8	17.8	901	20	US-10-739-930-3765	Sequence 3765, Ap
27	178.6	16.6	439	18	US-10-424-599-81721	Sequence 81721, A
28	175.2	16.2	1069	19	US-10-767-701-9317	Sequence 9317, Ap
29	169	15.7	777	21	US-10-738-922-1	Sequence 1, Appli
30	169	15.7	777	21	US-10-738-927-1	Sequence 1, Appli
31	156.6	14.5	777	21	US-10-738-922-2	Sequence 2, Appli
32	156.6	14.5	777	21	US-10-738-927-2	Sequence 2, Appli
33	154.6	14.3	1007	18	US-10-425-114-21611	Sequence 21611, A
34	151.6	14.1	1018	20	US-10-425-115-119846	Sequence 119846, A
35	146	13.5	1025	18	US-10-425-114-13774	Sequence 13774, A
36	146	13.5	1092	20	US-10-425-115-70617	Sequence 70617, A
37	145.6	13.5	1165	18	US-10-424-599-62644	Sequence 62644, A
38	133	12.3	1039	18	US-10-425-114-24225	Sequence 24225, A
39	133	12.3	1249	20	US-10-425-115-179879	Sequence 179879, A
40	131.8	12.2	1148	19	US-10-437-963-29636	Sequence 29636, A
41	128.2	11.9	996	19	US-10-767-701-15339	Sequence 15339, A
42	126.2	11.7	483	10	US-09-770-961-854	Sequence 854, App
43	125.2	11.6	1155	15	US-10-163-198-100	Sequence 100, App
44	120.6	11.2	1584	18	US-10-424-599-82031	Sequence 82031, A
45	110.6	10.3	1006	20	US-10-425-115-36145	Sequence 36145, A

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SABB2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; PRIOR FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-10-780-002-1

Query Match 100.0% Score 1079; DB 21; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 4.1e-295;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGCCGGGAAAGAAAGAACTTAAACAGGCATTAATCAATGAAGGAAACACT 60  
|||||







QY 61 TTGTTTGTAGTACATGGTGCATGCGAGGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
DB 209 TTGTCTGTGTCATGTATAGGCATAGTGCCTGGTGTGGTACAGCTTAAGCCACTGT 268  
QY 121 TAGAAGCTGCAGGCCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180  
DB 269 TGGAAATCCGCGGCCCAAAAGTCACAGTCTTGACCTTGCAGCTTCTGGCATCGACACAC 328  
QY 181 GAAATATAGAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTGATGGAGTTGATGG 240  
DB 329 ACGACATGGAAGACATCCACACATTTCTGAGTATTTCTAAGCCCTTTGTTGGATCTCTTGG 388  
QY 241 AATCTCTTTTTCAGCAGATGAGAAGGTTATATTAGTGGGCGATAGTCTTTGGTGGTATCAAT 300  
DB 389 CGTGGCTTCTCTAATGAAGAGTGGTCTTGTGGGCGATAGTCTTTGGAGGATCAGTA 448  
QY 301 TGGGACTTGTATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
DB 449 TAGCCCTTGCATGGACAAATCCAGAGAAATATCACTTGGAAATTTTCTTAAACAGCTT 508  
QY 361 TCATGCTGATTTCTGTTCAAACTCTCTCTTTGTTTGGACAGATATATGAGCGACGC 420  
DB 509 TTGTTCTGTATACCCCAACAAACCATCACTGCTTTAGAAGAGTACATTTGATAGATACC 568  
QY 421 CAGCCGAGAATTTGGTGGATACCTAGTCTTTTACCATATATGTTTCCCTGGAAGGCCACTGA 480  
DB 569 CATATACCGGATGGATGGACACTGAGCTCTGGATAGTGGAGC-----AAAA 616  
QY 481 CATCCATGTTTTTGGCCCAAGTTCTTGGCTCAAGCTCTACCAAGCTATGCTCTCCTG 540  
DB 617 CAACATTTGTTTTGGCATCAAAATTTCTGTCACCTAAGTTCTATCAACTCTGCTCCACTG 676  
QY 541 AGGATCTTCATTTAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTAT 600  
DB 677 AGGATCTGGAATTTGGTGAAGACTTTTAAAGAAAGAGGTTTCACTATTTGCTGGAAGACTTT 736  
QY 601 CGAAGCCCAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGCT 660  
DB 737 CTAAAGCAGAGAAATTTTCAAGAGAAAGATGGGTCTGTTTCCAGTGTCTTATTTATTT 796  
QY 661 GCATGAGGATTAAGGCATACCAAGAAATTTCCAGCGATGGCAAAATTTGACAACTTTGGTG 720  
DB 797 CCAATGAGGACTTGGTAATTTCCAAAGGAGTATCAGCAATGGATGATCCAAATGACAGGA 856  
QY 721 TCACTGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCAATGCTATCGGAGCCCAAA 780  
DB 857 TTGATGTGTGCGAGAGATCAAGGGATCAGATCAGATGATGTTTATGCTTTAGCAAAACCCACA 916  
QY 781 AACTTTGGCGCTCTCTCTTGGAAATTTGCCATAA 814  
DB 917 AACTATGTTTATCTCTCTCGAGATAGCTGATAA 950

## RESULT 7

US-10-780-002-33  
; Sequence 33, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-P02652W0  
; CURRENT APPLICATION NUMBER: US/10780,002  
; PRIORITY FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 33

; LENGTH: 870  
; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-33

Query Match 25.4%; Score 274.6; DB 21; Length 870;  
Best Local Similarity 60.4%; Pred. No. 6.7e-67;  
Matches 489; Conservative 0; Mismatch 314; Indels 6; Gaps 2;

QY 13 AAAAGAACTAACAGGCGATAAAAATTCAAATGAAGAGGAGAAACACCTTTGTTTTHAGTAC 72  
DB 1 AAAAAGTACGAAAGGAAAATATGAGTGAGGAGAGGAGGAAACACACTTCGCTAGTAC 60  
QY 73 ATGTTGTCATGCCATGAGGTTTCGAGTTTGTGTACAAGCTAAAGCCACTGCTAGAGCTGCAG 132  
DB 61 ATGTTGCTGCTCCAGCGCATGTTGCTGTGTACAGGTTAAGCTCTTCTTCGAGGCTTTGG 120  
QY 133 GCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGA---TTTGAGAAAAATAG 189  
DB 121 GCCATCGTGTAAACCGCTTAGACCTAGCTAGCTTCCGGTATAGACACACAGGTCAATCA 180  
QY 190 AGGAGCTTCGCACACTTTATGATTTATATCTTTGCCATTTGATGGAGTTGATGGAATCTCTTT 249  
DB 181 CTGACATTTCTACATGTGAACAATATTTCTGAGCCATTTGATGAGCTAATGATCTTATTGC 240  
QY 250 CAGCAGATGAGAGGTTATATTAGTGGGCGCATAGTCTTGTGGTATGAATTTGGGACTTG 309  
DB 241 CGAATGATGAGAGGTTGACTCTGTTGTCATAGCTTTGGAGGTTTGAAGTTAGCCCTTAG 300  
QY 310 CTATGAAAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCCCTG 369  
DB 301 CCATGGAATGTTTCCCGATAAATCTCTGCTCTGCTCTTCTGCTGCTGCTGCTGCTGCTG 360  
QY 370 ATTCGTTTCAAACTCTCTCTTTGTTTGGAAACAGTATAATGAGGAGGAGCCAGCGGAGA 429  
DB 361 ACACAAACACTCACCATCGTTTCTGCGAGGAAAGTTTGCAGAGCAGCATGACACCAAG 420  
QY 430 ATTCGTTTGGATCTCAGTTTTCATCATATGTTTCCCTGGAAGAGCCACTGACATCCATCT 489  
DB 421 GATGATGGGCTCTGAGCTCGAGATATATGTTT---CAGATTAATTCGGCTTGTCTGTGT 477  
QY 490 TTTTGGGCCCCAAGTTTCTTGGCTCAAGCTCTACAGCTATGCTCTCTCTGAGGATCTTG 549  
DB 478 TCTTCAGCACCGACTTCATGAAGCACCGTCTCTACCACTTCTCTCTGTTGGAGGATCTTG 537  
QY 550 CATTAGCATCATCATTTGTTGAGACCAAGCTCTTGTGTTTATGGAAGACCTATCGAAGGCCA 609  
DB 538 AGCTTGGATTTGCTTCTTAAAGAGGCTAGTTCTATTTTATTAATGAATTTATCGAAGATGG 597  
QY 610 AGTATTTTCCAGATGAAAGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGCTGCTGAGG 669  
DB 598 AGAATTTTTCGAGAAAGGTTATGATCTGTTCTCGAGCTTACATTTGTGTCGAAGAGG 657  
QY 670 ATAAAGGCATACCAAGAAATTTCCAGCGATGGCAAAATTTGACAACTTGTGTCTACTGAAG 729  
DB 658 ACAACATTTATCTCGAAGAGCCATCAACGATGGATCCATTAATTTATCCGGCGAATTTAG 717  
QY 730 CAATAGAGATTAAGGTTGCTGATCATGGAATGCTATGCGAGGCCCAAAACCTTTGGG 789  
DB 718 TGATTTGATGGAAGAGACTGATCATATGCAATTTTGTGCAAACTCTCACTACTAAGTG 777  
QY 790 CCTCTCTCTTGGAAATTTGCCATTAATATAC 818  
DB 778 ACCATCTATTGGCAATCGCTGACAAATTTTC 806

## RESULT 8

US-10-780-002-34  
; Sequence 34, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda

```
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF
; FILE REFERENCE: 3670-P02652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-34

Query Match      25.2%; Score 271.6; DB 21; Length 837;
Best Local Similarity 59.9%; Pred. No. 4.6e-66;
Matches 492; Conservative 0; Mismatches 324; Indels 6; Gaps 2;

QY      6 GCGAAGAAAGAACTCAACAGGCATAAATTCAAATGAAGGAAGGAAAAACACTTGT 65
DB      19 GAGACAAAGTTAAAGTAGAGAAATAGCTGAGGAGGAGGAAGCAACACGCTGT 78

QY      66 TTAGTACATGGTGCATGCCATCGAGGTGGAGTTGGTACAAAGCTAAAGCCACTGCTAGAA 125
DB      79 CTAGTACATGGTGTCTGCCATGGCGCTGGTCTGTGTAACAAGTTAAAGCCGAGCTGAG 138

QY      126 GCTGAGGCCCATAAAGTTACAGCCCTTGATTTAGCAGCTTCGGCACTGATTTG---AGA 182
DB      139 GCTTCTGGCCACCGGTAAACCGCCGTAGATCTAGCTGCTCCGGTATAGACATGACCAGG 198

QY      183 AAAATAGAGGAGCTTCGCACACTTTATGATTAATCTTTGCCATTCATGGAGTTGATGAA 242
DB      199 TCATCAGACATATATCCACATCGCGAACATFACTAGAGCCATTCATGACGTAATGACC 258

QY      243 TCTCTTTCAGCAGATGAGAAAGTTATATTAGTGGGCAATGCTTTGGTGGTATGAATTTG 302
DB      259 TCACCTACCATGATGATGAGAAGGTTGTGCTTGTGGTTCATAGCTTAGGAGGTTTGAGTTTA 318

QY      303 GCACTTGCTATGGAAAGTATCCACAAAGATCTATGCTGCTGTTTCTTGCTGCTCTTTC 362
DB      319 GCTATGGCCATGGATATGTTCCGACCAAAATCTCTGTTCTTGTGCTGCTGCTATG 378

QY      363 ATGCTCTGATCTGTTCACAACTCCTCTTGTGTTTGGAAACAGTATAATGAGCGGACGCCA 422
DB      379 ATGCCAGACACCAACACTCACCATCCTTCGTATGGATAAGCTAAGAAAAGAACTTCA 438

QY      423 GCCGAAATTTGGTGGATACCTCAGTTTTPAACATATGGTTCCCTCGAAGAGCCACTGACA 482
DB      439 CGAGAGGAATGGTTAGACACCGTGTTT---ACGAGCGAAGAACTGATTTTCTCAGCGAG 495

QY      483 TCATGTTTTTTGGGCCAAAGTTCTGGCTCAACAGCTCTACAGCTATGCTCTCTGTAG 542
DB      496 TTTTGGATTTTGGACCAAGATTCATGGCCAGAAGCTTGTATCAGTTGTTCCAGTCCAA 555

QY      543 GATCTTGCAATAGCATCATCTTGTGGACCAAGCTCTTCTGTTTATGGAAGACCTATCG 602
DB      556 GATCTTGAATTTGGCGAAGATTTGTTGAGGGCAACCCATTCATTAAGAAAGATATGGA 615

QY      603 AAGGCCAAGTATTTTCAACATGAAGCGTTTGGATCAGTGAAGAGAGTTTACATTTGTTGTC 662
DB      616 GAGAGAAGAGCTTCAGTGGAGGAAGGATACGGATCCGTTACACGTATATTTATTTGATGC 675

QY      663 ACTGAGGATAAAGGCATACAGAGNATTCAGCGATGCGAAATTCAGCAACATTTGGTGTTC 722
DB      676 GGAAGAGATCTTGTGTACCCGAAGATTTACCGCGATCGATGATCAGCAACTTCCCCCA 735

QY      723 ACTGAAGCAATAGAGATTAAGAGTGTCTGATTCATGCGCAATGCTATGCGAGCCCAAAAA 782
DB      736 AAGAAGTAAATGGAGATCAAGACGAGATCATATGCCAATGTTCTCCAAAGCTCAACAA 795
```

```
QY      783 CTTTGGCGCTCTCTCTGGAAATTTGCCCATAAATACAACTGA 824
DB      796 CTATGTGCTCTCTCTTGGAGATTGCAAAATAAATATGCTAA 837

RESULT 9
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551

Query Match      25.1%; Score 271; DB 9; Length 792;
Best Local Similarity 61.4%; Pred. No. 6.6e-66;
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY      55 AACACTTGTGTTTGTAGTACATGGTCATGCCATGGAGGTTGGAGTTGGTACAGCTAAAGC 114
DB      23 AACACTTGTGCTGTAGTACATGGTCGTGCCACGGCGCATGGTCTGTGACAAAGGTTAAGC 82

QY      115 CACTGCTAGAAAGCTGCGAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174
DB      83 CTCCTCTCGAGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTCTCCGTAAG 142

QY      175 A---TTTGAGAAAATAGAGGAGCTTCGCAACACTTTATGATATATCTTTGGCATGTAGG 231
DB      143 ACACAAACAGGTCAATCACTGACATTTCTACATGTGAACAAATATCTGAGCCATTTGATGC 202

QY      232 AGTTGATGGAATCTTTTACAGCAGATGAGAAGGTTATATTAGTGGGCGCATAGCTTTGGTG 291
DB      203 AGCTAATGACTTCATTTGCCGAATGATGAGAAGGTTGTACTCGTTGGTCATAGCTTTGGAG 262

QY      292 GTATGAATTTGGGACTTCTGCTATGAAAAGTATCCAAAAAGATCTATCTGCTGTTTTCT 351
DB      263 GTTTGAGTTTACGCTTAGCCATGATAGTTTCCGATAAATCTCTGCTCTGCTTTCG 322

QY      352 TGGCTGCTTTTATGCTGATTTCTGTTACAACTCTCTCTCTTTGTTTGGAAACAGTATAATG 411
DB      323 TGACTGCAATTCATGCCCGCACCAACAACTCACCATCGTTCTCGAGGAAAGTTTGCAG 382

QY      412 AGCGACCCGAGCCGAGAAATTTGTTGGATACCTAGTTTTTACCATATGTTTCCCTGAAG 471
DB      383 CGAGCATGACACCAAGAGGATGGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATA 439

QY      472 AGCCACTGACATCATGTTTTTTGGCCCAAAAGTTCTTTGGCTCACAAGCTCTACCAAGCTAT 531
DB      440 ATTCCGGCTTCTCTGTTCTTCAGCACCGACTTCATGAAGCACCGTCTCTACCAACTTT 499

QY      532 GCTCTCTCGAGATCTTGCATTTAGCATCATCAITTTGGTGAGCAAGCTCTTTGTTTATGG 591
DB      500 CTCCTGTGGAGATCTTGAGCTTGGATTTGCTTTTAAAGAGGCGCTAGTTTCATTTGTTATTA 559

QY      592 AAGACCTATCGAAGGCCAAGTATTTTACAGATGAACGGTTTGGATCACTGATGAAGAGAGTTT 651
```





292	GTATGAATTTGGGACTTGCTATGGAAAAAGTATCCAAAAAGATCTATGCTGCTGTTTCTT	351
293		
294		
295		
296		
297		
298		
299		
300		
301		
302		
303		
304		
305		
306		
307		
308		
309		
310		
311		
312		
313		
314		
315		
316		
317		
318		
319		
320		
321		
322		
323		
324		
325		
326		
327		
328		
329		
330		
331		
332		
333		
334		
335		
336		
337		
338		
339		
340		
341		
342		
343		
344		
345		
346		
347		
348		
349		
350		
351		
352		
353		
354		
355		
356		
357		
358		
359		
360		
361		
362		
363		
364		
365		
366		
367		
368		
369		
370		
371		
372		
373		
374		
375		
376		
377		
378		
379		
380		
381		
382		
383		
384		
385		
386		
387		
388		
389		
390		
391		
392		
393		
394		
395		
396		
397		
398		
399		
400		
401		
402		
403		
404		
405		
406		
407		
408		
409		
410		
411		
412		
413		
414		
415		
416		
417		
418		
419		
420		
421		
422		
423		
424		
425		
426		
427		
428		
429		
430		
431		
432		
433		
434		
435		
436		
437		
438		
439		
440		
441		
442		
443		
444		
445		
446		
447		
448		
449		
450		
451		
452		
453		
45		

RESULT 12

```

US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217

```

Query Match 24.9%; Score 269; DB 11; Length 792;  
Best Local Similarity 61.1%; Pred. No. 2.4e-65;  
Matches 472; Conservative 0; Mismatches 295; Indels 5; Gaps 2;

Qy	55	AAACACTTGTGTTT	TAGTACAT	TGTCATG	CCATG	CGAGTT	TGGAGTT	TGGTAC	AGCTA	AAAGC	114
D <sub>b</sub>	23	AAACAGTCGTTCT	TAGTACAT	TGTCATG	CGAGCTT	TGGCGCT	TGGTGTG	TGTACAG	CTTAAGC	82	

QY	115	CAC	TGCT	TAA	AGCTG	CGAC	GGCCAT	TAAG	GGTTAC	CAG	CCCTTG	ATT	TAG	CAG	CTTCTG	GCAC	TG	174				
DB	83	CGC	AGCT	CGAG	CGCTCT	TGGCC	CACCG	CGTAAC	CGCG	GTAGATCT	AGCT	CGCTCCG	GTA	TAG	142							
QY	175	ATT	TG	--	AG	AAAA	TATAG	AGAG	CTTCG	CACACT	TTATG	ATTAT	CTTTG	GCAT	TGATGG	231						
DB	143	AC	ATGA	CCAG	GTCA	ATCAC	AGATAT	TCC	CATG	CGGA	CAAT	CTC	CAG	AGCC	AT	TGATGC	202					
QY	232	AG	TTG	ATG	GA	ATCT	CTTT	CAG	CAG	ATG	AG	AGG	TAT	ATT	TAG	TGGGG	CATAGCTTTGGTG	291				
DB	203	AG	CTAA	TG	AC	CTCA	CTAC	CAG	ATG	TGA	GAAG	TGT	CTG	TTG	TGG	CTATAG	CTAGCTTTAGGAG	262				
QY	292	GT	ATGA	ATT	TGG	AGCT	TGCT	ATG	GA	AAAG	TATCC	CA	AAA	AGATCT	ATG	CTG	CTGTTTCT	351				
DB	263	GT	TTG	AG	TTT	TAG	CTATG	TGCC	ATG	ATATG	TTT	TCC	AG	CAAA	ATCT	CTG	TTTCTG	322				
QY	352	TGG	CTG	CTT	CAT	GC	TGAT	CTG	TTCA	CAACT	CTC	CTCT	TTG	TTT	TGGA	CAAG	TATAATG	411				
DB	323	TG	ACTG	CTAT	GAT	GC	CAG	CA	CAAA	CACT	CAC	CACT	CTT	CG	TAT	GG	TAAAGTAAGAA	382				
QY	412	AG	CGGA	CG	CCAG	CGAG	AAATT	TGG	TTGG	TATCT	CAG	TTTTT	TAC	CA	TATGG	TTCC	CTGGAAG	471				
DB	383	AAG	AAACT	TTCA	CG	AG	AGAA	TGG	TTAG	CA	CGT	GTTT	--	AC	GAG	CGAG	AAACCTGAT	439				
QY	472	AG	CCA	CTG	AC	AT	CCAT	G	TTTTT	TGG	CCCA	AAAG	TTCT	TGG	CTCA	CAAG	CTCTAC	531				
DB	440	TT	CTAG	CG	AG	TTT	TGG	AT	TTT	TGG	AC	CAG	AA	TT	CAT	GC	CAAG	AACTTGTATCAGTTGT	499			
QY	532	GCT	CT	CT	CAG	AG	ATCT	TGC	AT	TAG	CA	TCA	T	TGG	TGAG	CA	CAAG	CTCTTTGTTTATGG	591			
DB	500	CT	CAG	TTCCA	AG	ATCTT	GGA	TTGG	CG	AAAA	TTG	TGG	TGAG	GGCA	AA	CC	CA	TTGATTTAAGA	559			
QY	592	AAG	AC	CTAT	CG	AG	CCCA	AG	TATTT	C	AG	ATG	AA	CGG	TTT	TGG	AT	CGATGGAAGAGATTT	651			
DB	560	AG	ATAT	TGG	CAG	AG	AG	AAAG	CTT	CAG	TG	AG	GA	AG	GAT	CGG	TTAC	ACG	TATAT	619		
QY	652	AC	ATT	TG	TG	CA	CTG	AG	AG	AT	AA	GG	CA	TAC	AG	AG	AA	TTCC	AG	CGATGG	711	
DB	620	TT	ATT	TG	TAT	TG	CGGA	AA	GGATCT	TGT	G	T	GT	CA	CC	CG	AA	TTA	CC	AG	CGATCG	679
QY	712	AC	ATT	TG	TG	TG	CTCA	TG	AC	CAAT	TAG	AG	AT	TAA	AG	TG	CTG	AT	CAC	ATG	GC	771
DB	680	ACT	TTT	CCCC	CA	AA	AG	AG	AT	TG	AG	AT	CA	AA	GC	CA	AT	CA	TAT	AT	GC	739
QY	772	AG	CCCC	CA	AAAA	CTTT	TG	CG	CC	CTCT	CT	CTTT	TG	GA	AA	TT	TGCC	CA	TAA	AT	CA	824
DB	740	AG	CT	CA	CA	CA	CT	AT	G	CT	CTCT	CTTT	TG	GA	AT	TG	CA	AA	TAA	AT	TA	792

## RESULT 13

```

US-10-425-114-12800
; Sequence 12800, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12800
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701212327_FLI
US-10-425-114-12800

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 01:44:06 ; Search time 173.605 Seconds  
(without alignments)  
8595.849 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagcgtattctcaacaa.....ttttcggggcaactttcatc 912

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221.6	24.3	1078	3	US-08-981-256A-4
2	43	4.7	7218	1	US-08-232-463-14
3	38.6	4.2	1227	4	US-09-252-991A-13741
4	38.6	4.2	1473	4	US-09-252-991A-13404
5	36.2	4.0	640681	4	US-09-790-988-1
6	35.4	3.9	459	3	US-09-427-700-8
7	35.4	3.9	572	1	US-07-989-363-1
8	35.4	3.9	572	1	US-08-264-526-1
9	35.4	3.9	1442	1	US-08-152-483B-8
10	35.4	3.9	4009	1	US-08-152-483B-6
11	35.4	3.9	4294	1	US-08-152-483B-2
12	35.4	3.9	42094	4	US-09-949-016-16445
13	34.8	3.8	832	4	US-09-621-976-2813
14	34.8	3.8	175265	4	US-09-949-016-16089
15	34.4	3.8	601	4	US-09-949-016-57364
16	34.4	3.8	49378	4	US-09-949-016-13408
17	34.4	3.8	317366	4	US-09-949-016-16001
18	34	3.7	312474	4	US-09-949-016-17434
19	33.8	3.7	154023	4	US-09-949-016-17057
20	33.6	3.7	1141	4	US-09-806-708B-22
21	33.6	3.7	1750	3	US-08-949-770-1
22	33.6	3.7	26763	4	US-09-949-016-11984
23	33.6	3.7	26764	4	US-09-949-016-15566
24	33.4	3.7	1122	4	US-09-252-991A-11137
25	33.4	3.7	1260	4	US-09-252-991A-10925
26	33.2	3.6	601	4	US-09-949-016-46899
27	33.2	3.6	747	4	US-09-252-991A-3590

c 28	33.2	3.6	1191	4	US-09-252-991A-3692	Sequence 3692, Ap
c 29	33.2	3.6	1845	4	US-09-252-991A-3639	Sequence 3639, Ap
c 30	33.2	3.6	246444	4	US-09-949-016-13113	Sequence 13113, A
c 31	33	3.6	2525	3	US-09-342-648-1	Sequence 1, Appli
c 32	33	3.6	26354	4	US-09-949-016-12746	Sequence 12746, A
c 33	33	3.6	26355	4	US-09-949-016-12995	Sequence 12995, A
c 34	33	3.6	140925	4	US-09-949-016-11777	Sequence 11777, A
c 35	33	3.6	140982	4	US-09-949-016-16295	Sequence 16295, A
c 36	32.8	3.6	601	4	US-09-949-016-79644	Sequence 79644, A
c 37	32.8	3.6	807	4	US-09-270-767-13804	Sequence 13804, A
c 38	32.8	3.6	1500	1	US-08-476-008-6	Sequence 6, Appli
c 39	32.8	3.6	1500	1	US-08-306-063-6	Sequence 6, Appli
c 40	32.8	3.6	1500	1	US-08-833-485-6	Sequence 6, Appli
c 41	32.8	3.6	1500	3	US-09-137-440-6	Sequence 6, Appli
c 42	32.8	3.6	1500	5	PCT-US91-06148A-6	Sequence 6, Appli
c 43	32.8	3.6	1673	1	US-08-476-008-4	Sequence 4, Appli
c 44	32.8	3.6	1673	1	US-08-306-063-4	Sequence 4, Appli
c 45	32.8	3.6	1673	1	US-08-833-485-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:  
; APPLICANT: Meinhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981.256A  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob  
; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-OZ1112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1078 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-981-256A-4	Query Match	24.3%	Score 221.6;	DB 3;	Length 1078;
	Best Local Similarity	55.3%	Pred. No. 6.2e-64;		
	Matches 431;	Conservative 0;	Mismatches 349;	Indels 0;	Gaps 0;
QY	76	TCACCTTCGTGTAGTTCACACCGCTTATCATGGAGCCTGGATCTCGTACAGCTCAAGCC	135		
DB	55	TCATTTTCTTCTTATTTCATATGCGCGTGCATGGATTTGGCACAGCTCAAGCC	114		

```
QY 136 CTTCTCTGTAATCAGCGCGGCCACCGCGTTACTGCTCTCGAATCTCGCGCGCTCCGGATCGA 195
Db 115 CTTCTCTGAGGCACTTGGCGACAAGGTTACTCACTGGACCTTGCAGCAAGCGGTTGA 174
QY 196 CCACGACCAATCCAGCGCGTTGNAACCGTTCGACCAATACTCCAAACCGTTGATCGAAGAC 255
Db 175 CCCAAGGCCAATTTGAGGAGATTGGCTCATTTGATGAGTATTTGAAACCTTGTTCAGCGTT 234
QY 256 CCTCAATCTCTTCCAGAGAACGAAAGAGTAAATCTGTTGATTTGAGTTTCAGCTTCGGAGGAT 315
Db 235 CTTGAGGCACTCCCTCCAGGGAAAGGTGATTTCTGGTTCGCGAGAGCTGTGGAGGACT 294
QY 316 CAACATCGCTCTCGCGCGCGACATATTTCCCGCGAAGATAAGGTTCTTGTGTTCTCTCAA 375
Db 295 CAATATAGCAATTTGCTGTGATAAATACTGTGAAAAGATTGCAAGTGTGTTTTCACAA 354
QY 376 CGCCTTCTTCCCGGACACCAACCGTCTTCAAGTCTTGGACCAAGTATATGAGAT 435
Db 355 TTCAGTATTGCGACACCGACACTGCCCATCTTACGTCGTGGATAAGCTCATGGAGGT 414
QY 436 GCGTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGGACGATGAG 495
Db 415 GTTCCCGACTGGAAGACACACAGTATTTAGGTACACTAAAGATGGCAAGGAGATAAC 474
QY 496 TTTATGAGATGGGACCAAAATCATGAAGCAAGTCTTTACCAAAATTTGCCCATAGA 555
Db 475 TGGATTGAACTGGGCTTCAAGGCTTCTGAGGAAAAATTTATATACCTTTTGGCGTCTGA 534
QY 556 GGAATACGAGCTGCAAAATGTTGCATGAGCAAGGTCATTTTTCAGAGGATCTATC 615
Db 535 GGAATATGAACTGGCGAAGATGTTGCAAGGAAGGATCATTTTCAAAATATTTTATAGC 594
QY 616 AAGAAAGAAAATTTAGCGGAAAGGATATGTTTCGTTGCAACGAGTTTACGTAATGAG 675
Db 595 TAAGCGACCATTTCTTCACTAAGGAGTTACGATCGATTAGAAAATTTATGTGTGAC 654
QY 676 TAGTGAAGCAAAAGCCATCCCTGCGATTTTCAATTCGTTGATGATGATAATTTCAAGT 735
Db 655 CGACCAAGACGAAATATTTTACCTGAAATTTCAACTCTGGCAATAGAAAACTATAAAC 714
QY 736 CTCGAAAGCTACGAGATCGATGGCGGAGATCACATGTTGATGCTCCAAACCCCAAA 795
Db 715 AGACAAGGTTTAAAGTTCGAAGGTCGAGATCATAAATTCGAGCTTACAAAGACTAAGA 774
QY 796 ACTCTTTGACTCTCTCTGCTATTGCCCACCGATTATATGTAATAATCTTAAGTCCGTT 855
Db 775 GATCGTGAATTTCTCCAGAGGTGGCTGATACCTATAATTGACTTCTTTGAGGCTTTT 834
```

## RESULT 2

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14
```

```
Query Match 4.7%; Score 43; DB 1; Length 7218;
Best Local Similarity 3.5%; Pred. No. 0.0065;
Matches 10; Conservative 165; Mismatches 110; Indels 0; Gaps 0;
```

```
QY 418 GGACAAGTATATGAGATGCTCGAGGTTTGGGAGATTTGAGTTTTCATCTCATGAAC 477
Db 1334 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1275
QY 478 AAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGCACGCTTTTA 537
Db 1274 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1215
QY 538 CCAAAATTTGCCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATT 597
Db 1214 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1155
QY 598 TTTCCACAGAGATCTATCAAGAAAGAAAGTTTTCGCGAAGAGATATGTTTCGTTGCA 657
Db 1154 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1095
QY 658 ACAGATTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGA 702
Db 1094 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1050
```

## RESULT 3

```
US-09-252-991A-13741/c
; Sequence 13741, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13741
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13741
```

```
Query Match      4.2%; Score 38.6; DB 4; Length 1227;
Best Local Similarity 54.6%; Pred. No. 0.056;
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 208 CCAGGCGGTGAACCGTCGACGAATACTCCAACCGTTGATCGAAACCCCTCAATCTCT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 CCAGCGGTGGCCCGGATCGACCACTACCGGACCTGTTGCTCGACACCCCTCAAGCGGC 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 268 TCCAGAGAAGAAAGAGTAAATTCGTGTTGATTGAGTTTCGGAGGATCAACATCGCTCT 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 CAGTGACGTGATCGCGGAGCGGTGTTGATTCGCGCGGACGTTTCAACAGCGCCTA 323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 328 CGCCGCCGACATATTTCCCGC 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CTTCGAACACGCTTCTCTGC 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-252-991A-13404
; Sequence 13404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13404
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13404

Query Match      4.2%; Score 38.6; DB 4; Length 1473;
Best Local Similarity 54.8%; Pred. No. 0.064;
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 208 CCAGGCGGTGAACCGTCGACGAATACTCCAACCGTTGATCGAAACCCCTCAATCTCT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 621 CCAGCGGTGGCCCGGATCGACCACTACCGGACCTGTTGCTCGACACCCCTCAAGCGGC 680
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 268 TCAGAGAACGAAGAGTAAATTCGTGTTGATTGAGTTTCGGAGGATCAACATCGCTCT 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 681 CAGTGACGTGMAATCGCGGAGCGTGTGTTGATTCGCGCGGACGTTTCAACAGCGCCTA 740
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 328 CGCCGCCGACATATTTCCCGC 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 CTTCGAACACGCTTCTCTGC 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match      4.0%; Score 36.2; DB 4; Length 640681;
Best Local Similarity 55.0%; Pred. No. 32;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 540 AAAATTGTCCTATAGAGGATTACGAGCTGGCGAAATAATTGTCATAGGCAAGGTCATTTT 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59751 AAAATTAAATTAATAATAGAAAGAACTGGCTTAACTTGAACAAGTAAAAAGATATTA 59692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 600 TCACAGAGGATCTATCAAAAGAAAGAAAGTTTATAGCGAGGAAGATATGTTCCGTCGAAC 659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59691 TAAAGAATGCTCTATTGGGAAGAAAGCTAGACGCGGAAAAAGAAATGTCGAAC 59632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 660 GAGTTTACG 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59631 TAAATTTACG 59623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-427-700-8/c
; Sequence 8, Application US/09427700
; Patent No. 6372489
; GENERAL INFORMATION:
; APPLICANT: Anticancer
; APPLICANT: Zhao, Ming
; TITLE OF INVENTION: METHOD AND MODEL FOR HAIR PIGMENTATION
; FILE REFERENCE: 31276-20017.00
; CURRENT APPLICATION NUMBER: US/09/427,700
; CURRENT FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: US 60/105,725
; PRIOR FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(450)
; OTHER INFORMATION: ORF 438 gene
US-09-427-700-8

Query Match      3.9%; Score 35.4; DB 3; Length 459;
Best Local Similarity 51.6%; Pred. No. 0.34;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 74 CATCACTTCGTGTAGTTTCACACCGCTTATCATATGAGCGCTGATCTGGTACAGTCAAG 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 CACCGGCTCGTAGTGGCTGACGACGCTGATCCACGAGCGCTGCGGCTTCCGATCACA 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 CCCCTCCTTGAATCAGCGCGCCACCGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTG 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 CAGTTGAGCGCGCTCCACGACACGCGCGTAAACCGCGCGCGCTGATGCGCCCGTGACC 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 GACCCACGACCAATCCAGGCGGTTGAAACCGTCGACG 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GTCACGCGCGTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-07-989-363-1/c
; Sequence 1, Application US/07989363
; Patent No. 5385841
; GENERAL INFORMATION:
; APPLICANT: Ortega, Jose Daza
; APPLICANT: Gill, Jose Antonio
; APPLICANT: Garcia, Tomas Vigal
```





```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/152,483B
/ FILING DATE: No. 5529909ember 12, 1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 7/857,602
/ FILING DATE: March 30, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 923,692
/ FILING DATE: July 31, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 600,244
/ FILING DATE: October 22, 1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 641,617
/ FILING DATE: January 16, 1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 737,899
/ FILING DATE: July 26, 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Halluin, Albert P.
/ REGISTRATION NUMBER: 25,227
/ REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 854-3660
/ TELEFAX: (415) 854-3694
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1442
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ DESCRIPTION: DNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM:
/ IMMEDIATE SOURCE:
/ CLONE:
/ FEATURE:
/
/ US-08-152-483B-8
/
/ Query Match 3.9%; Score 35.4; DB 1; Length 1442;
/ Best Local Similarity 51.8%; Pred. No. 0.77; 76; Indels 0; Gaps 0;
/ Matches 81; Conservative 0; Mismatches 0;
/
/ QY 74 CATCACTTCGTGTAGTTACAAACGCTTATCATGGAGCGCTGATCTGGTACAAAGCTCAAG 133
/ Db 542 CACCGGCTCGTAGTGTGCTGACGACGCTGATCCACGAGCGCTGGCGGTTGGGCATCATG 483
/
/ QY 134 CCCCTCCTTTGAATCAGCGGCCACCGGTTACTGTGTGAACTCGCGCCTCCGGGATC 193
/ Db 482 CAGTTCGAGCGCGTCACCAACGCGGTAACCGCGCGTGTGATGCGCGCGCGTGACC 423
/
/ QY 194 GACCCAGCAACATCAGCGCGGTTGAAACCGTGCAGC 230
/ Db 422 GTCACCGCGGTGACCGCGCGGTGGTGGCCCCCGCG 386
/
/ RESULT 10
/ US-08-152-483B-6/c
/ Sequence 6, Application US/08152483B
/ Patent No. 5529909
/ GENERAL INFORMATION:
/ APPLICANT: Della-Cioppa, Guy
/ APPLICANT: Kumagai, Monto
/ TITLE OF INVENTION: TYROSINASE-ACTIVATOR
/ TITLE OF INVENTION: PROTEIN FUSION ENZYME
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/
/
/ STREET: 2730 Sand Hill Road
/ CITY: Menlo Park
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 94025
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0,
/ SOFTWARE: Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/152,483B
/ FILING DATE: No. 5529909ember 12, 1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 7/857,602
/ FILING DATE: March 30, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 923,692
/ FILING DATE: July 31, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 600,244
/ FILING DATE: October 22, 1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 641,617
/ FILING DATE: January 16, 1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 737,899
/ FILING DATE: July 26, 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Halluin, Albert P.
/ REGISTRATION NUMBER: 25,227
/ REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 854-3660
/ TELEFAX: (415) 854-3694
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4009
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA
/ DESCRIPTION: NO
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM:
/ IMMEDIATE SOURCE:
/ CLONE:
/ FEATURE:
/
/ US-08-152-483B-6
/
/ Query Match 3.9%; Score 35.4; DB 1; Length 4009;
/ Best Local Similarity 51.6%; Pred. No. 1.6; 76; Indels 0; Gaps 0;
/ Matches 81; Conservative 0; Mismatches 0;
/
/ QY 74 CATCACTTCGTGTAGTTACAAACGCTTATCATGGAGCGCTGATCTGGTACAAAGCTCAAG 133
/ Db 2780 CACCGGCTCGTAGTGTGCTGACGACGCTGATCCACGAGCGCTGGCGGTTGGGCATCATG 2721
/
/ QY 134 CCCCTCCTTTGAATCAGCGGCCACCGGTTACTGTGTGAACTCGCGCCTCCGGGATC 193
/ Db 2720 CAGTTCGAGCGCGTCCACGAAACACGCGCGTAACCGCGCGTGTGATGCGCGCGCGTGACC 2661
/
/ QY 194 GACCCAGCAACCAATCCAGSCCGTTGAAACCGTGCAGC 230
/ Db 2660 GTCACCGCGGTGACCGCGCGGTGGTGGCCCCCGCG 2624
/
/ RESULT 11
```

US-08-152-483B-2/c  
; Sequence 2, Application US/08152483B  
; Patent No. 5529909  
; GENERAL INFORMATION:  
; APPLICANT: Della-Cioppa, Guy  
; APPLICANT: Kumagai, Monto  
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR  
; TITLE OF INVENTION: PROTEIN FUSION ENZYME  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 2730 Sand Hill Road  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0,  
; SOFTWARE: Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/152,483B  
; FILING DATE: No. 5529909ember 12, 1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 7/857,602  
; FILING DATE: March 30, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 923,692  
; FILING DATE: July 31, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 600,244  
; FILING DATE: October 22, 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 641,617  
; FILING DATE: January 16, 1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 737,899  
; FILING DATE: July 26, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 854-3660  
; TELEFAX: (415) 854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4294  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: DNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM:  
; IMMEDIATE SOURCE:  
; CLONE:  
; FEATURE:  
; US-08-152-483B-2  
Query Match 3.9%; Score 35.4; DB 1; Length 4294;  
Best Local Similarity 51.6%; Pred. No. 1.7;  
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 74 CATCACTTCGTGTAGTTCACAAACGCTTATCATGGAGCCCTGGATCTGGTACAAAGCTCAAG 133  
Db 3590 CACCGGCTGCTAGTGGCTGACGACGCTGATCCACGAGCGGTGCGGCTTGGCATCACATG 3531

QY 134 CCCCTCCTTGAATCAGCGCGCCACCGGTTACTCTGTGCAACTCGCGCCCTCCGGATC 193  
Db 3530 CAGTTTCGACGCCGTCCACGAAACACGCGGTAAACCCCGCGGTGATGGCGCGCGGTGACC 3471  
QY 194 GACCCAGACCAATCCAGGCGGTTGAAACCGTCGACG 230  
Db 3470 GTCACCGCGGTGACCGCGCGCGGTGTCGCGCCCGCCCG 3434  
RESULT 12  
US-09-949-016-16445  
; Sequence 16445, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16445  
; LENGTH: 42094  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1) .. (42094)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16445

Query Match 3.9%; Score 35.4; DB 4; Length 42094;  
Best Local Similarity 50.9%; Pred. No. 8.6;  
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
QY 398 CACGTGCTTCTCAGCTTCGACAAGTATATGAGATGCTCGGAGTTTGGGAGATTGT 457  
Db 10870 CATGTAGCTTTTAAAGTTTAGTTTAAATTTAAAGTAAATAATTAAGCCTTAGGCAACATAGC 10929  
QY 458 GAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGAAGATGGACCAAAA 517  
Db 10930 AAGACCTCATCTCTGAAAAAAGAAAAAAGATAAAGATTAGCTGACATGGTGTGTTACA 10989  
QY 518 TTCATGAAGGACGCTTTTACCAAAATTTGCCATAGAGATTAC 562  
Db 10990 TGCCTGTAGTCCAGCTACTCAGGAGGCTGAGATGGGAGGATTAC 11034

RESULT 13  
US-09-621-976-2813/c  
; Sequence 2813, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET. 054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832  
; TYPE: DNA

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

Query Match      3.8%; Score 34.8; DB 4; Length 832;
Best Local Similarity 10.1%; Pred. No. 0.82;
Matches 33; Conservative 148; Mismatches 145; Indels 0; Gaps 0;

QY 532 TCTTTACCAAAATGTCATGAGGATTCAGAGCTGGCAAAATGTTCCATAGGCAAGG 591
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 YYYWTKTWYTTTWTBMMWKKARRWYWKSTYACASRYRYRTWGWMMYWKMM 303
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
592 GTCATTTTTCACAGAGCTATCAAGAAAGAAAGTTTACGAGGAAGGATATGGTTC 651
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
302 STRWYTCMKWCWYKGRCAWTTMARGRWSTAWGKWSKMSMSMCTRMYYKKGSTYW 243
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
652 GGTGCAACGAGTTTACGTAATAGTAGTAGTGAAGCAAGCCATCCCTCGGATTTTCAT 711
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
242 TMKCTCATWYWKYKRWKMSKTCWSGSGGYMTSYSTRSYMYWASWMTMCMWGR 183
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
712 TTGGATGATGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCAT 771
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182 WMSYWMAGKWKWRYATTWRRAMMMWMAWMTMMYMWMMWCMSSRGAAMRYRTMMWMM 123
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
772 GGTGATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCGCACCGATT 831
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
122 YRYWKKSYRTRCAAWAWKTKRSYYWCRWKKRCMMMMMMMAWYKTMWMAACWKT 63
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
832 TATGTAATAATCTTAAGTCGGTTTA 857
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
62 RYRWAWAWRWMMWMMWMMYWYWR 37

RESULT 14
US-09-949-016-16089
; Sequence 16089, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16089
; LENGTH: 175265
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(175265)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16089

Query Match      3.8%; Score 34.8; DB 4; Length 175265;
Best Local Similarity 58.8%; Pred. No. 38;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 795 AACTCTTTGACTCTCTCTGCTATTGCGCAGGATATATGTAATAATCTTAAGTCGGTT 854
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
118139 AAGTGTGTTTTAAATACCTCTTCGAGTCTTACATATATGTAATAATCTTAAGCCGTT 118198
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
855 TTACTTTTTTCTCATCGTTTACTAATAAAACAAACCCCTTTT 896

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 118199
; US-09-949-016-57364/c
; Sequence 57364, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57364
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-57364

Query Match      3.8%; Score 34.4; DB 4; Length 601;
Best Local Similarity 59.0%; Pred. No. 0.89;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 4 ACGCTATTCTCAAAACAAGCAAGAACCTTAACCAAAAGAAACGCTCTACGAGACATAT 63
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 ACTCAAGTCTCAAAAAAAGAAATAATAATATTCAGATTCAAAAGATATAT 178
QY   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
64 GGAGAGGAAACATCATCTTCGTGTAGTTTCAACAACGCTTAT 103
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
177 GCCTAAGAATGTTTCATTACAGCATTTGTTTACAAATGATTTT 138

Search completed: September 12, 2005, 06:38:21
Job time : 177.605 secs
```

**This Page Blank (uspro)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 04:46:01 ; Search time 2020.05 Seconds  
(without alignments)  
2964.938 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagcgtattctcaacaa.....ttttccgggcaactttcattc 912

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	912	100.0	912	21	US-10-780-002-36
2	481.4	52.8	483	10	US-09-770-961-854
3	243	26.6	989	21	US-10-780-002-32
4	241.6	26.5	792	9	US-09-938-842A-1320
5	241.6	26.5	792	11	US-09-938-842A-1320
6	232.6	25.5	1079	21	US-10-780-002-1
7	223.4	24.5	933	18	US-10-424-599-64748

8	222.6	24.4	837	21	US-10-780-002-34	Sequence 34, Appl
9	218.8	24.0	792	9	US-09-938-842A-217	Sequence 217, App
10	218.8	24.0	792	11	US-09-938-842A-217	Sequence 217, App
11	214.2	23.5	792	9	US-09-938-842A-1551	Sequence 1551, Ap
12	214.2	23.5	792	11	US-09-938-842A-1551	Sequence 1551, Ap
13	214.2	23.5	870	21	US-10-780-002-33	Sequence 33, Appl
14	205.4	22.5	777	21	US-10-738-922-1	Sequence 1, Appli
15	205.4	22.5	777	21	US-10-738-922-1	Sequence 1, Appli
16	204.8	22.5	1101	19	US-10-437-963-62048	Sequence 62048, A
17	202.6	22.2	1158	18	US-10-424-599-85023	Sequence 85023, A
18	199.6	21.9	792	21	US-10-780-002-35	Sequence 35, Appl
19	198.6	21.8	777	21	US-10-738-922-2	Sequence 2, Appli
20	198.6	21.8	777	21	US-10-738-922-2	Sequence 2, Appli
21	194.4	21.3	771	9	US-09-938-842A-1278	Sequence 1278, Ap
22	194.4	21.3	771	11	US-09-938-842A-1278	Sequence 1278, Ap
23	194.4	21.3	771	21	US-10-780-002-40	Sequence 40, Appl
24	192.2	21.1	834	21	US-10-780-002-39	Sequence 39, Appl
25	191.6	21.0	1035	18	US-10-425-114-12800	Sequence 12800, A
26	185.2	20.3	783	21	US-10-780-002-38	Sequence 38, Appl
27	176.8	19.4	819	21	US-10-780-002-37	Sequence 37, Appl
28	172.6	18.9	901	21	US-10-780-002-41	Sequence 41, Appl
29	171.2	18.8	1069	19	US-10-767-701-9317	Sequence 9317, Ap
30	149.6	16.4	996	19	US-10-767-701-15339	Sequence 15339, A
31	146.4	16.1	653	18	US-10-424-599-52897	Sequence 52897, A
32	145.2	15.9	1148	19	US-10-437-963-29636	Sequence 29636, A
33	141.2	15.5	786	19	US-10-437-963-62697	Sequence 62697, A
34	134.8	14.8	1173	19	US-10-437-963-89948	Sequence 89948, A
35	132	14.5	1007	18	US-10-425-114-21611	Sequence 21611, A
36	132	14.5	1018	20	US-10-425-115-119846	Sequence 119846, A
37	128.8	14.1	1025	18	US-10-425-114-13774	Sequence 13774, A
38	128.8	14.1	1092	20	US-10-425-115-70617	Sequence 70617, A
39	126	13.8	818	18	US-10-425-114-16669	Sequence 16669, A
40	126	13.8	830	20	US-10-425-115-156630	Sequence 156630, A
41	118.4	13.0	1233	19	US-10-437-963-19868	Sequence 19868, A
42	118.2	13.0	920	18	US-10-424-599-128189	Sequence 128189, A
43	117.2	12.9	439	18	US-10-424-599-81721	Sequence 81721, A
44	117	12.8	901	20	US-10-739-930-3765	Sequence 3765, Ap
45	116	12.7	907	18	US-10-425-114-24341	Sequence 24341, A

#### ALIGNMENTS

##### RESULT 1

US-10-780-002-36

; Sequence 36, Application US/10780002

; Publication No. US20050034196A1

; GENERAL INFORMATION:

; APPLICANT: Kleisig, Daniel F.

; APPLICANT: Kumar, Dhirenda

; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN

; FILE REFERENCE: 3670-PO2652W0

; CURRENT APPLICATION NUMBER: US/10/780,002

; CURRENT FILING DATE: 2004-02-17

; PRIOR APPLICATION NUMBER: PCT/US02/26312

; PRIOR FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 60/312,863

; PRIOR FILING DATE: 2001-08-16

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 36

; LENGTH: 912

; TYPE: DNA

; ORGANISM: Arabidopsis Thaliana

US-10-780-002-36

Query Match 100.0%; Score 912; DB 21; Length 912;

Best Local Similarity 100.0%; Pred. No. 4.7e-278; Mismatches 0; Indels 0; Gaps 0;

1 ATCAGC GTATTCTCAACAAAGCAAGACCTTAAACAAAGAAACGCTTACGAGACA 60

|||||



Publication No. US20050034196A1  
GENERAL INFORMATION:  
APPLICANT: Klessig, Daniel P.  
APPLICANT: Kumar, Dhirendra  
TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
FILE REFERENCE: 3670-PO2652WO  
CURRENT APPLICATION NUMBER: US/10/780,002  
CURRENT FILING DATE: 2004-02-17  
PRIOR APPLICATION NUMBER: PCT/US02/26312  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: 60/312,863  
PRIOR FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 32  
LENGTH: 989  
TYPE: DNA  
ORGANISM: Arabidopsis Thaliana  
US-10-780-002-32

Query Match 26.6%; Score 243; DB 21; Length 989;  
Best Local Similarity 59.6%; Pred. No. 9.7e-66;  
Matches 468; Conservative 0; Mismatches 305; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCACTTCGTGTAGTTTCAACGCTTATCATGAGCGCTCGATCTGGTACA 125  
Db 120 AGAGGAAACATCACTTCGTGTAGTTTCAACGCTTATCATGAGCGCTCGATCTGGTACA 179  
QY 126 AGCTCAAGCCCTCCTTGAATCAGCGCGCACCGCTTACTGTGTGAACTCGCGCT 185  
Db 180 AGGTTAAGCCGCTGTAGAGCGGTGGCCACCGCTACTGTGTGAACTTAGCTGCT 239  
QY 186 CCGGATCGACCA----CGACCAATCCAGCGCTTGAACCGTCCGAGCAATCTCCAAAC 242  
Db 240 CCGGAATAGACACACAGAGTCGATCACTGACATCCCACTCGCAACAACTACGAGC 299  
QY 243 CGTTGATCGAAACCTCAATCTTCCAGAGAACGAGAGTAATCTGTTGGATTCA 302  
Db 300 CATTTGACGAGCTCTTGACTCTATTGCGCAATGATGAAGGTTGTCTGTTGGTACA 359  
QY 303 GCTTCGAGGATCAACATCGCTCTCGCGCGCACATATTTCCGCGCAAGATTAAAGTTTC 362  
Db 360 GCTTTGCTGCTTGAACCTTAGCATAGCCATGAAAGTTTCCCGAATAAATCTCTGTG 419  
QY 363 TTGTGTTCTCAAGCCTTCTTCCCGCACACCAACCGTCCCTTCTCAAGTTCTGGACA 422  
Db 420 CTGTATTTCTGACTGCTTTTCATGCGCGCACACCGAACACTCAACATCTCTGTTGGACA 479  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGAGATGTGAGTTTCTCATGTGAA 476  
Db 480 AGTTTGGAAACAACTGCTCAAGAAGCATGATGGGCAACCGAAATTCGAACCTTATGTT 539  
QY 477 CAAGAAATGGAGATGAGTTTATTGAAGATGGACCAAAATTCATGAAGCGACGCTCTTT 536  
Db 540 CAGACAAATTCGAGCTGAGTATGTT---TTTTAGCCTGACTTCAAGATTGGGTCTCT 596  
QY 537 ACCAAATTTGCCATAGAGGATACGAGCTGGCAAAAATTTGTCATAGGCAAGGTCAT 596  
Db 597 ACCAGCTTTCTCCAGTTGAGGATCTTGAATCTGGATTACTTTTAAATGAGCCAGGATCT 656  
QY 597 TTTTCACAGGATCTCAAGAAAGAAAGTTTAGCGAGGAGGATGTTGGTTC 656  
Db 657 TATTTAATTAACGATTTATCGAAGATGAAAAATCTTCGGATGAAGGATGTTGTTCT 716  
QY 657 AACGAGTTTACGTAATGAGTAGGAAGACAAAGCCATCCCTGCGATTTCAATCGTTGGA 716  
Db 717 CTCGAGTTTTCATGTTGTGAAGAGACAAAGCAATTCAGAAAGACCGCCAGAGATGA 776  
QY 717 TGATGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCAATGGTGA 776  
Db 777 TGATGATAATTTTCGCGTGAATTTTAGTGATGAGATGAGGAGACAGATCATATGCCAA 836

QY 777 TGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCGATTATGT 836  
Db 837 TGTTCTGCAAGCCTCAGCACTCAGTATTACTTCTGAAATCGGACCAATTCGTTT 896  
QY 837 AATAA 841  
Db 897 AATCA 901  
RESULT 4  
US-09-938-842A-1320  
Sequence 1320, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
FILE REFERENCE: SRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 1320  
LENGTH: 792  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1320

Query Match 26.5%; Score 241.6; DB 9; Length 792;  
Best Local Similarity 59.6%; Pred. No. 2.4e-65;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCACTTCGTGTAGTTTCAACGCTTATCATGAGCGCTGGATCTGGTACA 125  
Db 14 AGAGGAAACAACTTTCTACTAGTACATGTTCTGTGCAATGCGCGCTGGTCTGTGACA 73  
QY 126 AGCTCAAGCCCTCCTTGAATCAGCGCGCACCGGTTACTGCTGCTGCACTCGCGCT 185  
Db 74 AGTTAAGCCGCTGCTAGAGCGGTGGGCGCACCGGTAACCTGCTGTGACTTAGTGTCT 133  
QY 186 CCGGATCGACCCA---CGACCAATCCAGGCGGTTGAAACCGTCCGCAATACTCCAAAC 242  
Db 134 CCGGATAGACACACAGAGGTCGATCACTGACATCCCACTGCGAACATACTCGGAGC 193  
QY 243 CGTTGATCGAAACCTTCAATCTCTTCCAGAGAACGAGAGTAATCTGTTGGATTCA 302  
Db 194 CATTGACGAAGCTCTGACCTCATTTGCCAAATGATGAAGAGTTGTCTGTTGGTACA 253  
QY 303 GCTTCGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGCGCAAGATTAAAGTTTC 362  
Db 254 GCTTTGTTGGCTTGAACCTTAGCCATAGCCATAGGAAAGTTTCCCGAATAAATCTCTGTG 313  
QY 363 TTGTGTTCTCAAGCCCTTCTTGGCCGACACAAACCGCTGCTTCTCACGTTCTGGACA 422  
Db 314 CTGTATTTCTGACTGCTTTTCATGCGGACACCGAACACTCAACATCTTCTGTTGGACA 373  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA 476  
Db 374 AGTTTGAAGCAACATGCTCAAGAAGCATGATGGGCAACCGAAATTCGAACCTTATGTT 433  
QY 477 CAAGAAATGGAGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTT 536  
Db 434 CAGACAAATTCGCGAGTGAATGTT---TTTTAGCCCTGACTTCAATGAAGTTGGGTCTCT 490  
QY 537 ACCAAATTTGTCCTCCATAGGATTACGAGCTGGGCAAAAATGTTGATAGGCAAGGTCAT 596

Db 491 ACCAGCTTCTCCAGTTGAGGATCTTGAACTGGGATTAATCTTTAATGAGGCGAGATCGT 550  
Qy 597 TTTTTCACAGAGGATCTATCAAAAGAAAGAAAGTTTATAGCGAGGAGATATGGTTGCGTGC 656  
Db 551 TATTTATTAAAGATTTATCGAAGATGAAAGAACTTCTCGGATGAAGATATGGTCTGTTC 610  
Qy 657 AACGAGTTTACGTAATGATAGTAGTGAAGACAAAGCCATCCCTGCGCATTTCAATCGTTGGA 716  
Db 611 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAAAGAACGCCAGAGATGGA 670  
Qy 717 TGATTGATAAATTTCAACGCTCTCGAAGTCTAGAGATCGATGGCGGAGATCACATGGGA 776  
Db 671 TGATTGATAAATTTCCCGTGAAATTTAGTAGGATGGAGATGAGGACAGATCATATGCCAA 730  
Qy 777 TGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTTGCCACCGATTATATGT 836  
Db 731 TGTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAATCGCGACAAATTCGTTT 790  
Qy 837 AA 838  
Db 791 AA 792

## RESULT 5

US-09-938-842A-1320  
; Sequence 1320, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1320  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1320

Query Match 26.5%; Score 241.6; DB 11; Length 792;  
Best Local Similarity 59.6%; Pred. No. 2.4e-65;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
Qy 66 AGAGAAACATCATTCTGGTTAGTTTCAACAGCTTATCATGGAGCGTGGATCTGGTACA 125  
Db 14 AGAGAAACAAATTTTGTACTAGTACATGTTTCTGTCATGGCGTGGTCTGTGTTACA 73  
Qy 126 AGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTCGAACCTCGCGGCT 185  
Db 74 AGGTTAAGCCGTGTAGAGGCGGTGGCCACCGGTAATCTGCTGTGGACTTAGCTGCTCT 133  
Qy 186 CCGGATCCACCA---CGACCAATCCAGGCGGTTGAAACCGTGCAGCAATACTCCAAAC 242  
Db 134 CCGGATAGACACAGGAGTCTGATCACTGAATCCCCACATCGCGACATTAATCTCGGAGC 193  
Qy 243 CGTTGATCGAAACCTTCAAAATCTTTCCAGAGAAAGAGAGTAAATCTTGGTTGGATTCA 302  
Db 194 CATTGACGAAGCTCTGACTCATTTGCCAAATGATGAAAGGTTTGTGCTCGTTGGTTCACA 253  
Qy 303 GCTTCGAGGATCAACATCGTCTCGCGCGGACATATTTTCGCGGAGATTAAAGTTTC 362

Db 254 GCCTTGGTGGCTTGAACTTAGCCATAGCCATGGAAGATTTCCCGAAAAAATCTCTCTGTCG 313  
Qy 363 TTGTGTTCTCAACGCCCTTCTTGCCGACACCAACCCAGTGCCTTCTCAAGTTCTTGACA 422  
Db 314 CTGTATTTCTTGACTGCTTTTATGCGCGACACCGAACACATCCATCTCTGCTTGGACA 373  
Qy 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATTGTGAGTATTTTCTCTCATGAAA 476  
Db 374 AGTTTGAAGACATGCTCAAGACATGGATGGGACCGAATTCGAACCTTATGGTT 433  
Qy 477 CAAGAAATGGGACATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGCACATCTTT 536  
Db 434 CAGACAAATCCGGACTGAGTATGTT---TTTtagccctgacttcattgaaagtgggtctct 490  
Qy 537 ACCAAATTTGCCATAGAGGATTTACGAGCTGGGCAAAATTTTCATAGGCAAGGTCAT 596  
Db 491 ACCAGCTTTCTCCAGTTGAGGATCTTGAACCTGGGATTAATTTTAAATGAGGCGAGATCGT 550  
Qy 597 TTTTTCACAGAGATCTATCAAGAAAGAAAGTTTtagcgagaaaggatattggttcggtgc 656  
Db 551 TATTTTAAAGATTTTTCGAAGATGAAGAACTTCTCGGATGAAGGATATGGTCTGTTC 610  
Qy 657 AACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCAATCGTTGGA 716  
Db 611 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAAAGAACGCCAGAGATGGA 670  
Qy 717 TGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGA 776  
Db 671 TGATTGATAATTTTCCGGTGAATTTAGTAGTGGAGATGGAGGAGACAGATCATATGCCAA 730  
Qy 777 TGCTCTCAACCCCAAAACTCTTGAATCTCTCTCTGCTTATTTGCCACCGATTATATGT 836  
Db 731 TGTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAATTCGCGACAAATTCGTTT 790  
Qy 837 AA 838  
Db 791 AA 792

## RESULT 6

US-10-780-002-1  
; Sequence 1, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-10-780-002-1

Query Match 25.5%; Score 232.6; DB 21; Length 1079;  
Best Local Similarity 56.3%; Pred. No. 2e-62;  
Matches 458; Conservative 0; Mismatches 349; Indels 6; Gaps 1;  
Qy 33 AACCAAAAGAAACGTTCTACGGACATATGGAGAGAAACATCACTTCGTGTAGTTC 92  
Db 13 AAAAGAAACTAACAGGCAATAAAATTCAAATGAAGGAAGGAAACACATTTGTTTAGTAC 72  
Qy 93 ACAAGCTTATCTAGGAGCTCGGTACAAAGCTCAAGCCCTCTCTGAATCAGCGG 152



Db 73 ATGGTCATGCCATGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGCTAGAAAGTCGAC 132  
QY 153 GCACCGCGTTACTGCTCGAACTCGCGCTCCGGGATCGACCCACGACCAATCCAGG 212  
Db 133 GCATTAAGGTTACAGCCCTTGATTAGCAGCTTCTGGCACTGATTTGAGAAAATAGAGG 192  
QY 213 CGTTGAAACCGTCGACGAATACCTCAAACCGTTGATCGAAACCCCTCAAATCTCTTCCAG 272  
Db 193 AGCTTCGCACACTTTATGATGATTAACCTTGCCTTGAATGAGTTGATGGAATCTCTTCAG 252  
QY 273 AGAACGAAGAGGTAATCTGTTGGATTGAGTTCCGAGGCGATCAACATCGCTCTCGCG 332  
Db 253 CAGATGAGAAGGTTATATAGTGGGCATAGTCTTGGTGTATGAATTTGGGACTTGTCTA 312  
QY 333 CCACATATTTCCGGCGAAGATTAAAGTTCTTGTTCTCTCAACGCTTCTTGCCCGACA 392  
Db 313 TGGAAAGTATCCAAAGAGTCTATGCTGCTGTTTCTTGCTGCTTTCATGCTGATTT 372  
QY 393 CAACCCAGTGCCTTCTCAGTTCTGCAAGATPATATGAGATGCTCGAG-----GTT 446  
Db 373 CTGTTCAACAACCTCTCTCTTGTGTTGGAACAGTATTAATGAGCGGACGCCAGCCGAGAA 432  
QY 447 TGGAGATTGTGAGTTTTCATCTCATGAACCAAGAAATGGGAGGATGAGTTTATGAAGA 506  
Db 433 GGTGGATCTCAGTTTTCACCATATGTTTCCCTGAAGAGCCACTGACATCCATGTTT 492  
QY 507 TGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCATAGAGGATTACGAGC 566  
Db 493 TTGGCCCAAAGTTCTTGGCTCAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCA 552  
QY 567 TGGCAAAATTTGTCATAGCGAAGGTCATTTTTCACAGAGGATCTATCAAAAGAAAGAA 626  
Db 553 TAGCATCATATTTGGTGAGACCAAGCTCTTTGTTATGGAAGACCTATCGAAGGCCAAGT 612  
QY 627 AGTTTACGAGGAGAGATATGTTTGGTGCACAGGATTTACGTAATGATGATGAGACA 686  
Db 613 ATTTACAGATGAACGTTTGGATCAGTGAAGAGGTTTACATTTGTGTCACTGAGGATA 672  
QY 687 AAGCCATCCCTGCGATTTTCATTCGTTGGATGATGATTAATTTCAACGTCCTGAAAGTCT 746  
Db 673 AAGGCATACAGAGAAATTCAGCGATGGCAATTTGCAACATTTGGTGTCACTGAAGCNA 732  
QY 747 ACGAGATCGATGGCGAGATCACATGGTGTGCTCTCCAAACCCCAAAACCTCTTTGACT 806  
Db 733 TAGAGATTAAAGTGCTGATCAGTGGCAATGCTATGCGAGCCCAAAACCTTTGCGCT 792  
QY 807 CTCTCTCTGATATGCCACCGATTATATGTAAT 839  
Db 793 CTCTCTTGGAAATTCGCCATAAATACAACTGAT 825

RESULT 7

US-10-424-599-64748  
; Sequence 64748, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 64748  
; LENGTH: 933  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29480C.1  
US-10-424-599-64748

Query Match 24.5%; Score 223.4; DB 18; Length 933;  
Best Local Similarity 55.9%; Pred. No. 1.6e-59;  
Matches 447; Conservative 0; Mismatches 346; Indels 6; Gaps 1;

QY 75 ATCACTTCTGTTAGTTTACAAAGCTTATCATCGAGCCCTGGATCTGGTACAAAGCTCAAGC 134  
Db 44 AGCACTTTGTTCTGTTGATGGGATGCCATGGAGCATGGTGTGGTATTAAGCTCAAGC 103  
QY 135 CCTCTTGAATACGCGCCACCGCTTACTGCTGTCGAACCTCCGCGCTCCGGGATCG 194  
Db 104 CAGGTTTGAATCTGCTGGGCACAGGTCACAGCTTGACCTTGACGCTTCTGAGCCA 163  
QY 195 ACCACGACCAATCCAGCCGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAAA 254  
Db 164 ACATGAAGAAATTTGAAGATGTTGACATTTTTCACAGTACTGAGCCTTTGTTGTTTC 223  
QY 255 CCTCAATCTCTCCAGAGAACGAGGTAATTTCTGGTTGGATTCAGCTTCGAGGCA 314  
Db 224 TATTGGACACAATTCCTCAATGAAAGGATGTTTCTAGTTGGTCCACAGCTTTGGAGGC 283  
QY 315 TCAACATCGCTCTCGCGCCGACATATTTTCGGGCGAAGATTAAAGTCTCTGTGTTCTCA 374  
Db 284 TGAACATAGCACTTCCCATGGAGAAATTTTCAGAAAAGGTAGCAGTTGGTGTCTTCA 343  
QY 375 ACGCTTCTTGCCCGACACAAACCCACGTCCTTCTCAGCTTCTGACGAAGTATATGGAGA 434  
Db 344 CAGCTTTTGTCTCCAGACGTTGNAACACACCCATCTTATGCTCTTGGAAAGTACAGCGAGA 403  
QY 435 TGCCTGAGGTTTGGGAGATTTGAGTGTTCATCTCATGAAACAAGAAATGGGACGATGA 494  
Db 404 GACCCCGCTAGCTGATGCTAGTACACTGAATTTGCTCCAAGTGGAAACAAACAA 463  
QY 495 GTTTATTGAAGTGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCCATAG 554  
Db 464 TGTCTTT-----GGCCCCAACTTCTTGTCCGCAAGCTCTACCAACTATCCCCAATG 517  
QY 555 AGGATTACGAGCTGGCAAAATTTTGCATAGCGAAGGCTCATTTTTCACAGAGGATCTAT 614  
Db 518 AGGATTTTGAATTTGGCAAGACTTTTAAGGCCATCATCACTCTTCATGGAAGACTTGA 577  
QY 615 CAAGAAAGAAAGTTTACGAGGAGGATATGTTTGGTGGCAACGAGTTTACGTAATGA 674  
Db 578 TAAACAAAGAACTTCTCAAAGAGGATATGGTCACTGCGACGCTTATTTGTTT 637  
QY 675 GTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTTCAAG 734  
Db 638 GCACTGAGGACCTTGCATTTCCATTTGGAATATCAGCTCTTCATGATCCAAATTTGGGT 697  
QY 735 TCTCGAAAGTCTACGAGATCGATGGCGAGATCACATGGTGTGATCTCTCCAAACCCCAAA 794  
Db 698 TCAATGAAGTTGTAGAGATCAAAGGCGAGACCATATGGCTATGCTTTCGAAGCCCAAG 757  
QY 795 AACTCTTGTACTCTCTCTGCTATGTCGATGTCACCGATATATGTAATTAATCTTAAAGTCGTT 854  
Db 758 AACTATTGATTTCCCTCCAGCAGATAGCGACTAAATATGCAATTAATTCACGATGCT 817  
QY 855 TTACTTTTCTCATCGTT 873  
Db 818 TTAATTAGTTTATTTGTT 836

RESULT 8

US-10-780-002-34  
; Sequence 34, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002

```
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-34
```

```
Query Match      24.4%; Score 222.6; DB 21; Length 837;
Best Local Similarity 56.0%; Pred. No. 2.6e-59;
Matches 464; Conservative 0; Mismatches 359; Indels 6; Gaps 2;

QY 16 AACAGCAAGAACCTTAACCAAAAGAAACGCTACGGAGACATATGGAGAGGAACA 75
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9 AAGAAGTCGAGAGACAAAGTTAAAAGTAGAAGAAATGAGTAGAGGAGGAGGAACA 68
QY 76 TCATCTCGTGTAGTTCAACACGCTTATCATGGAGCCTGGATCTCGGTACAAGCTCAAGCC 135
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 ACACGTCGTTCTAGTACATGGTGCTTGCATGCGCCTGGTCTGTTACAGGTTAAGCC 128
QY 136 CTTCTTGAATCAGCGGCGCACCGGTATCTGCTGTGAACTGCGCGCTCCGGATCGA 195
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 GCAGCTCGAGGCTTCTGGCCACCGGCTAACCCGCTAGATCTAGCTGCGCTATAGA 188
QY 196 C---CCACGACCAATCCAGGCGGTTGAAACCGTCCAGCAATACTCCAACCGTTGATCGA 252
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 CATGACCAAGGTCAATCACAGATATATCCACATGCGCAACAATACTCAGAGCCATGATCGA 248
QY 253 AACCCCTCAAAATCTTCCAGAGAACGAGAGTAATTTCTGTTGATTCAGCTTCGGAGG 312
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 GCTAATGACCTCACTACACAGATGATGAGAAGTTGTGCTTGTGTCATAGCTTAGGAGG 308
QY 313 CATCAACATCGCTCTCGCGCGCGACATATTTCCGGCGAAGATTAAGGTTCTGTGTTCT 372
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
309 TTTGAGTTTGTAGTATGGCCATGGATATGTTTCCGACCAAAATCTCTGTTCTGTCTTGT 368
QY 373 CAAACGCTTCTGCGCGCACACAAACCGTCTCAGCTTCTCAGCTTCGGAACAAGTATGGA 432
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 GACTGTATGATGCCAGACACCAACATCACCATCTCTGATGCGGATAGCTAAGAAA 428
QY 433 GATCGCTGAGGTTTGGGAGATTGTGA---GTTTTATCTCATGTAAGAAACAAAGAAATGGGAC 489
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 AGAAACTTCACGAGAGGAATGTTAGACACCGTGTTTACGCGGAGAAACCTGATTTTCC 488
QY 490 GATGAGTTTATGAGATGGGACCAAAATTCATGAGGACACGCTCTTTACCAAAATTTGCC 549
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 TAGCGAGTTTGGATTTTGGACCAAGAAATTCATGCGCAAGAACTTGTATCAGTTGTCTCC 548
QY 550 CATAGAGGATACGAGCTGGCAAAATGTTGATAGGCAAGGTCATTTTTCACAGAGA 609
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
549 AGTCCAAGATCTTGAATTTGGCGAAATGTTGTGAGGCGAAACCCCATGATTAAGAAAGA 608
QY 610 TCTATCAAAAGAAAGAAAGTTTAGCGGAGGAGATATGTTGCGTGCACAGATTACGT 669
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
609 TATGCGAGAGAGAGAGCTTCAGTGAGGAGGATACGGATCCGTTACAGTATATTTAT 668
QY 670 AATGAGTAGTGAAGCAAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATGATTAATTT 729
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 TGTATGCGGAAAGGATCTTGTGTACCCCGAAGATTACCAGCGATCGATCAGCAACTT 728
QY 730 CAAAGTCTGAAAGTCTACGAGATGATGGCGGAGATCACATGGTGATGCTCTCCAAACC 789
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 TCCCCCAAAGAGTAATGGAGATCAAGCCGAGATCATATGCCAAATGTTCTCCAGCC 788
QY 790 CCAAAACCTCTTTGACTCTCTCTGCTATTGCCACCGGATTAATGTA 838
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 TCAACAACATATGTGCTCTCTCTGAGATTGCAAAATAATATGCTAA 837
```

```
RESULT 9
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217

Query Match      24.0%; Score 218.8; DB 9; Length 792;
Best Local Similarity 56.6%; Pred. No. 4.1e-58;
Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2;

QY 55 GAGACATATGAGAGGAGAAACATCACTTCGTGTTAGTTTCAACACGCTTATCATGAGCGCTG 114
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 GAGTCAGGAGGAGGAGCAACACGTCGTTCTAGTACATGTCGTCCTTCCCATGCGCCTG 62
QY 115 GATCTGTTACAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGCTGTCGA 174
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 GTGCTGGTACAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGCTTAACCGCGGTAGA 122
QY 175 ACTCGCGCCTCCGGGATCGAC---CCACGACCAATCCAGCCGTTGAAACCGTCGACGA 231
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 TCTAGTCGCTCCGTTATAGACATGACCCAGGTCAATCACAGATATATCCATGCGAACA 182
QY 232 ATACTCCAAACGTTGATCGAAACCTCAATCTCTCCAGAGAACGAGGTAATTTCT 291
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 ATACTCAGAGCAATTGATGACGCTAATGACCTCTACAGATGATGAGAAGTTGTGCT 242
QY 292 GGTGGATTCAAGTTTCGAGGACATCAACATCGCTCTCGCCCGCACATATTTCCGCGCAA 351
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 TGTGGTCATAGCTTAGGAGGTTTGTAGTTAGCTATGCGCATGATATGTTTCCGACCA 302
QY 352 GATTAAGGTTCTTGTTTCTCAACGCTTCTTCCCGACACAAACCCACGTCGCTTCTCA 411
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 AATCTCTGTTCTGCTTTGTGACTGTCTATGATGCCAGACACCAAACTCACCATCCTT 362
QY 412 CGTTCGACAAAGTATATGAGGATCGCTCGAGGTTTGGAGATTGTGA---GTTTTCATC 468
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 CGTATGGGTAAGCTTAAGAAAGAAACTTCAAGAGAGGAATGGTTAGACACCGTGTGTAC 422
QY 469 TCATGAACAAGAAATGGGACGATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGC 528
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
423 GAGCGAGAAACCTGATTTTCTAGGAGTTTGTGATTTTGGACCAAGATTCATGGCCAA 482
QY 529 ACGTCTTTTACCAAAATTTGCCCATAGAGGATTAAGAGCTGGCAAAATTTGTCATAGCA 588
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 GAACTGTATCAGTTGTCTCCAGTCCAAAGATCTTGAATTTGGCGAAATTTGTTGGTGAGGC 542
QY 589 AGGGTCATTTTTCAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGG 648
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
543 AAACCCATTGATTAGAAAGATATGGCAGAGAGAGAGCTTTCAGTGGAGGAGGATACGG 602
QY 649 TTCGTCGACAGGTTTACGTAATAGTAGTAGTGAAGACAAAGCCATCCCTCGATTTTCAT 708
```

Db 603 ATCCGTTACAGTATATTATTATGATGCGAAGGATCTTGTTGTCACCCGAAAGATTACCA 662  
Qy 709 TCGTTGGATGATTAATTTTCAACGCTCTCGAAGTCTACGAGATCGATGGCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCCGCAAAAGAAAGTAATGGAGATCAAAAGACGAGATCA 722  
Qy 769 CATGGTGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCGCACCGA 828  
Db 723 TATGCCAATGTTCTCAAGCCTCAACACTATGCTCTCTCTTGGAGATTGCAATAA 782  
Qy 829 TTATATGTAA 838  
Db 783 ATATGCTAA 792

RESULT 10  
US-09-938-842A-217  
; Sequence 217, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 217  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-217

Query Match 24.0%; Score 218.8; DB 11; Length 792;  
Best Local Similarity 56.6%; Pred. No. 4.1e-58;  
Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2;  
Qy 55 GAGACATATGGAGGAAACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCCTG 114  
Db 3 GAGTGAGGAGGAGGAGGAAAGCAACACGCTGTTCTAGTACATGTTGCTTGCATGGCGCCTG 62  
Qy 115 GATCTGGTACAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCA 174  
Db 63 GTGCTGGTACAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGGTTAACCCGCTAGA 122  
Qy 175 ACTCGCGGCTCCGGATCGAC---CCACGACCAATCCAGGCGGTTGAAACCGTCGACGA 231  
Db 123 TCTAGCTGCTCCGTTATAGACATGACAGGTTCAATCAGATATATCCAGATGCGAACA 182  
Qy 232 ATACTCCAAACCGTTGATCGAAACCCCTCAAACTCTCTTCAGAGAACGAGAGTAATCT 291  
Db 183 ATACTCAGAGCCATTGATCGAGCTAATGACCTCACTTACCAGATGATGAGAAGGTTGTGCT 242  
Qy 292 GGTGTGATTCAGTTTCGGAGGATCAACATCGCTCTCGCGCGCGGATATTTCCGCGCA 351  
Db 243 TGTGTGTCATAGCTTAGGAGGTTTGAAGTTTAGCTATGGCCATGGATATGTTTTCCGACCA 302  
Qy 352 GATTAAGGTTCTTGTTCTCAAGCCCTTCTTGGCCGACACAAACCGATGCTCTTCA 411  
Db 303 AATCTCTGTTCTGTTCTTGTGACTGTAAGTGCAGACACCAACACTCAACATCTCTT 362  
Qy 412 CGTCTTGGAACAAGTATATGAGATGCTCGAGGTTTGGGAGATTGTGA---GTTTTCATC 468

Db 363 CGTATGGGTAAGCTAAGAAAAAGAAACTTTCAGAGAGGAATGGTTAGACACCGTGTTCAC 422  
Qy 469 TCATGAAACAGAAATGGACGATGAGTTTATTGAGATGGGACCAAAATTCATGAAGGC 528  
Db 423 GAGCGAGAAACCTGATTTTCTAGCGAGTTTGGGATTTTGGACCCAGAAATTCATGGCCAA 482  
Qy 529 AGCTCTTTTACCAAAATTTGTCATAGAGATTCAGAGCTGGCAAAATGTTTCATAGGCA 588  
Db 483 GAACTTGATGATGTTCTCCAGTCCAAGATCTTGAATTTGGCGAAAAATGTTGGTGGGCG 542  
Qy 589 AGGGTCATTTTTTCAGAGGATCTATCAAAAGAAAGAAAGTTTGGAGGAAAGGATATGG 648  
Db 543 AAACCCATTGATTAAAGAAAGATATGGCAGAGAGAAAGAGCTTCAGTGGAGAAAGGATACG 602  
Qy 649 TTCGGTGCAACGAGTTTACGTAATGAGTAGTAGAACAAGCCATCCCTGCGATTTCAT 708  
Db 603 ATCCGTTACACGTATATTATTGATGCGGAAGGATCTTGTGTACCCGGAAGATTACCA 662  
Qy 709 TCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCCGCAAAAGAAAGTAAATGGAGATCAAAAGACGAGATCA 722  
Qy 769 CATGTGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGGCCACCGA 828  
Db 723 TATGCCAATGTTCTCCAAAGCCTCAACAACATATGCTCTCTCTTGGAGATTGCAATAA 782  
Qy 829 TTATATGTAA 838  
Db 783 ATATGCTAA 792

RESULT 11  
US-09-938-842A-1551  
; Sequence 1551, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1551  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1551

Query Match 23.5%; Score 214.2; DB 9; Length 792;  
Best Local Similarity 57.3%; Pred. No. 1.2e-56;  
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;  
Qy 55 GAGACATATGGAGGAAACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCCTG 114  
Db 3 GAGTGAGGAGAGGAAAGCAACACTTCGTGTAGTACATGGTGGTGGCGGCGATG 62  
Qy 115 GATCTGGTACAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGCTGTGCA 174  
Db 63 GTGCTGGTACAGGTTAAGCCTCTTCTCGAGGCTTTGGGCCATCGTTAAACCGCCTTAGA 122  
Qy 175 ACTCGCGCCTCCGGGATCGACCCA---CGCAAAATCCAGGCGGTTGAAACCGTTCGACGA 231  
Db 123 CCTAGCTGCTTCCGGTATAGACAAACCGAGGTCAATCACTGACATTTCTTACATGTGAACA 182

```
QY 232 ATACTCCAAACCGTTGATCGAAACCTCMAATCTCTTCAGAGAACGAGGTAATTTCT 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ATATTCTGAGCCATTGATGACGCTAATGACTTTCAATGGCGAATGATGAGAAGTTGTACT 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGGATTACGCTTCGAGGAGCATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAA 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CGTTGGTCATAGCTTTGGAGGTTTGAGTTAGCTTTAGCCATGGATAGTTTCCCGATAA 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTTGTTCTCAACGCTCTTTCGCCGACACAAACCCACGTCCTTTCTCA 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AATCTCTGTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CTTCTGAGCAAGTATP-----ATGAGATGCTCGAGGTTTGGAGATTTGAGTTTTC 465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CGTCGAGGAAAAGTTTGCAGAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGA 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAAACAAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAA 525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GACATATGGTTCAGATAATTCGGGCTTGTCTGTGTT---CTTCAGACCGACTTCATGAA 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCAGCTCTTTACCAAAATGTCATAGAGGATTCAGAGCTGGCAAAATGTTGCATAG 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCACCGTCTTACCAACTTTCTCTCTGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 539
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTAGGAGGAGGATA 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 GCCTAGTTCATTTTATTAATGAATATTCGAAGATGGAGAACTTTTCTGAGAAAAGGTA 599
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTGGTGCAACGAGTTTACGTAATGATGATGAGTTCAGAGCTGGCAAAATGTTGCATAG 705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 TGAATCTGTTCTCGAGCTTACATTTGTTGTCGAAGAGGCAACATATCTCGGAAGACCA 659
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCGTTGGATGATGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGA 765
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 TCAACGATGGATGATCATAATTTATCCGGCGAATTTAGTGAATGAGATGGAAGAGACTGA 719
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGATGCTCTCAAAACCCCAAAACCTTTTGACTCTCTCTCTCTCTCTCTCTCT 825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 TCATATGCCAATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGA 779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGATT 830
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 CAATT 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551
```

```
Query Match 23.5%; Score 214.2; DB 11; Length 792;
Best Local Similarity 57.3%; Pred. No. 1.2e-56;
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;

QY 55 GAGACATATGAGAGGAAACATCACTTCGTGTTAGTTTCAACAACGCTTATCATGAGGCTG 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GAGTGAGGAGAGGAGCAACACTTCGTGCTAGTACATGTCGTCGTCGCCACGCGCATG 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 GATCTGGTAAAGCTCAAGCCCTTCCTTGAATAGCCGCGCCACCGCTTACTGTCTGCA 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 GTGCTGGTACAAAGGTTAAGCCTCTTCTCGAGGCTTTGGGCCATCGTGTAAACCGCTTAGA 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ACTCGCGCTCCGCGATCGAACCA---CGACCAATCCAGCCGCTTGAACCGCTCGACGA 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 CCTAGCTGCTTCCGGTATAGACACACAGGTCATCACTGACATTTCTCATGTGAACA 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 ATACTCCAAACCGTTGATCGAAACCTCMAATCTCTTCAGAGAACGAGGTAATTTCT 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ATATTCTGAGCCATTGATGACGCTAATGACTTTCAATGGCGAATGATGAGAAGTTGTACT 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGGATTACGCTTCGAGGAGCATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAA 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CGTTGGTTCATAGCTTTGGAGGTTTGAGTTAGCTTTAGCCATGGATAGTTTCCCGATAA 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTTGTTCTCAACGCTCTTTCGCCGACACAAACCCACGTCCTTTCTCA 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AATCTCTGTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CGTTCTGACAAAGTATP-----ATGAGATGCTCGAGGTTTGGAGATTTGAGTTTTC 465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CGTCGAGGAAAAGTTTGCAGAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGA 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAAACAAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAA 525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GACATATGGTTCAGATAATTTATTAATGAATATTCGAAGATGGAGAACTTTTCTGAGAAAAGGTA 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCAGCTCTTTACCAAAATGTCATAGAGGATTCAGAGCTGGCAAAATGTTGCATAG 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCACCGTCTTACCAACTTTCTCTCTGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 539
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTAGGAGGAGGATA 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 GCCTAGTTCATTTTATTAATGAATATTCGAAGATGGAGAACTTTTCTGAGAAAAGGTA 599
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTGGTGCAACGAGTTTACGTAATGATGATGAGTTCAGAGCTGGCAAAATGTTGCATAG 705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 TGAATCTGTTCTCGAGCTTACATTTGTTGTCGAAGAGGCAACATATCTCGGAAGACCA 659
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCGTTGGATGATGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGA 765
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 TCAACGATGGATGATCATAATTTATCCGGCGAATTTAGTGAATGAGATGGAAGAGACTGA 719
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGATGCTCTCAAAACCCCAAAACCTTTTGACTCTCTCTCTCTCTCTCTCTCT 825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 TCATATGCCAATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGA 779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGATT 830
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 CAATT 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-780-002-33
; Sequence 33, Application US/10780002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: ENCODING NUCLEIC ACIDS, SABB2, AND METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
```

```

; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-33

Query Match      23.5%; Score 214.2; DB 21; Length 870;
Best Local Similarity 57.3%; Pred. No. 1.2e-56;
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;

QY 55 GAGACATATGGAGAGAAACATCATCTGTTAGTTTCAACAGCCTTATCATGGAGCCTG 114
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 23 GAGTGAAGGAAGGAAGCAACACTTCGTGCTAGTACATGGTGGCGCCGCGCATG 82
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 GATCTGTTCAAGCTCAAGCCCTCTCTTGAATCAGCGCGCCACCGGTTACTGCTGCA 174
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 GTGCTGGTACAGGTTAAGCCTCTTCTGAGGCTTTGGGCCATCGGTAAACGCCCTTAGA 142
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ACTCGCGCCTCCGGATCGAACCA---CGACCAATCCAGCGCGTTTGAACCGTCGACGA 231
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 CCTAGCTGCTTCGGGTATAGACAAACAGGTCAATCACTGACATTTCTACATGTGAACA 202
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 ATACTCAAACCGTTGATCGAAACCGTCAAACTCTTTCAGAGAACGAAGAGTAATCT 291
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 203 ATATTCGAGCAATGATGAGCTAATGACTTCAITGGCCGAATGATGAGAAGGTTGTACT 262
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGTGATTCAGCTTCGGAGGATCAACATCGCTCTCGCGCGCGACATATTTCCGCGAA 351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 263 CGTGTGTCATAGCTTTGGAGGTTTGAGTTAGCCTTAGCCATGATGATGATTTCCCGATAA 322
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTTGTGTTCTCAACGCTTTTCCCGCGACACACACCGCTGCTTCTCA 411
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 323 AATCTCTGCTCTGCTCTGCTGACTGATTCATTCATGCGCGGACACCAACACTCACCACGTT 382
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CGTTCGGAACAAGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATGTGAGTTTC 465
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 CGTCGAGGAAAGTTTGCAGGACGATGACACCAAGGATGGATGGGCTCTGAGCTCGA 442
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAACAAGAAATGGGACGATGATTTATTGAAGATGGGACCAAAATTCATGAA 525
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GACATATGGTTAGATAATTCGGGCTTGCTGTGTT---CTTCAGCACCACCTTCATGAA 499
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCACGCTTTTACCAAAATTTGCCATAGAGGATTACGAGCTGGCAAAAATGTTGCATAG 585
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 500 GCACCGCTCTACCAACTTCTCTCTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 559
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGGTCATTTTTCACAGAGATCTATCAAGAAGAAAGTTTATAGCGAGAGAGATA 645
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 560 GCCTAGTTTCACTTTTATTAATGAATATCGAAGATGGAGAACTTTTCTGAGAAAGGTA 619
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTCGGTGCACGAGTTTACGTAATGATGATGAGTGAAGCAAGCCATCCCTGCGATT 705
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 620 TGGATCTGTTCTCGAGCTTACATGTTGTGCAAGAGGACCAATATTCGGAAGACCA 679
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCCTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGA 765
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 680 TCAACGATGATGATCCATAATTTATCCGGGAATTTAGTGAATGATGAGAGAGACTGA 739
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGTGCTCTCOAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTGGCCAC 825
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 740 TCATATGCCAAATGTTTTGCAAAACCTCAACTACTTAAGTGACCCTACTATTGGCAATCGCTGA 799
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGAATT 830
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 800 CAATT 804
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 14
US-10-738-922-1
; Sequence 1, Application US/10738922
; Publication No. US20050032159A1
; GENERAL INFORMATION:
; APPLICANT: Ichige, Eita
; APPLICANT: Semba, Hisashi
; APPLICANT: Imanaka, Tadayuki
; TITLE OF INVENTION: METHOD FOR LOW-TEMPERATURE CULTURE OF
; TITLE OF INVENTION: MICROORGANISM
; FILE REFERENCE: 08917-091001
; CURRENT APPLICATION NUMBER: US/10/738,922
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: JP 2002-365676
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Manihot esculenta Crantz
US-10-738-922-1

Query Match      22.5%; Score 205.4; DB 21; Length 777;
Best Local Similarity 55.3%; Pred. No. 7.2e-54;
Matches 420; Conservative 0; Mismatches 336; Indels 3; Gaps 1;

QY 77 CACTTCGTGTTAGTTTCAAAACGCTTATCATGAGGCTTGATCTGGTACAGCTCAAGCCC 136
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 13 CATTTTGTGTTGATTCATACCATTTTGCCATGCTGATGGATTTGGCATTAAGCTCAAAACA 72
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 CTCCTTGAATCAGCCGGCCACCGCGTTACTGCTGTCGAACCTCCGCGCTCCGGGATCGAC 196
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 GCCCTTGAGAGCTGGCCACAAAGTCACTGCACTGCACTGCACTGCACTGCACTGCACTG 132
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 CCACGACCAATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAAACC 256
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 133 CCAAGGCAAAATGAGCAGATTAATTCATTTGATGAATACTCTGAAACCCCTTATTGACTTTC 192
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 257 CTCAAATCTCTCCAGAGAACGAAGAGTAACTCTGTTGGATTGATTCAGCTTCGGAGGCATC 316
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 TTGGAGAAATCCCTCAAGGGGAAAGGTCATCATTTGTTGGAGAGCTGTGAGGGCTG 252
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 317 AACATCGCTCTCGCGCGACATATTTCCGCGGAAAGATTAAGGTTCTTGTGTTCTCTCAAC 376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 253 AATATTGCTATTGCTGCTGATAGATACGTTGACAAATTTGCAGCTGGTGTGTTCCAAAT 312
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 377 GCCTTCTTCCGCGACACACCCACCGTCTCTACGTTCTGGAACAAGTATATGAGATG 436
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 TCCTTTATTTGCCAGACACCGTTTCATAGCCCATCTTACACTGTGGAAGAGCTTTTGGAGTCG 372
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 437 CCGTGGAGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGG---GACGATG 493
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 TTTCTGATCGAGAGACACAGAGTATTTTACGTTTCTCAATATATCATCTGGAGAGACAAT 432
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 AGTTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCATA 553
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 ACAACATGAGCTGGGCTTCGTACTCTCTGAGGGAATTTATTTTACCAATGCACTGAT 492
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 GAGGATTCAGAGCTGGCAAAAATTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTA 613
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 GGGGAATATGAATCGCAAAAATGTAATGAGGAAGGATCACTGTTTCAAAATGTTTGG 552
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 TCAAGAAGAAAGTTTAGCGGAAGGATATGTTTGGTTCGTCGACAGGATTTACGTAATG 673
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 553 GCTCAGAGCCGAAGTTTCCGGAAGAGTTTACCGGATCAATTAAGAAAGTTTATATTGG 612
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 674 AGTAGTGAAGCAAAAGCCATCCCTGCGATTTTCACTTCTGTTGATGATTAATTTCAAC 733
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 613 ACCGATCAAGACAAATATTTTACCAGACTTTTCAACGCTGGCAATTTGCAAACTACAAA 672
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 GTCTCGAAGTCTACGAGATCGATGGCGGAGATCATGTTGGTGAATCTCTTCCAAAACCCCAA 793
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 673 CCAGACAAGGTTTATCAGGTTCAAGGTGGAGATCATAGCTCCAGCTTACAAAACTGAG 732

Qy 794 AAACCTCTTTGACTCTCTCTCTGCTATTTGCCACCGATTAT 832

Db 733 GAGGTAGCTCATATTCTCCAGAGGTGGCTGATGCATAT 771

RESULT 15

US-10-738-927-1

; Sequence 1, Application US/10738927

; Publication No. US20050032191A1

; GENERAL INFORMATION:

; APPLICANT: Semba, Hisashi

; APPLICANT: Ichige, Eita

; APPLICANT: Mukoyama, Masaharu

; TITLE OF INVENTION: METHOD FOR PRODUCTION OF

; TITLE OF INVENTION: S-HYDROXYNITRILE LYASE BY USE OF ESCHERICHIA COLI

; FILE REFERENCE: 08917-092001

; CURRENT APPLICATION NUMBER: US/10/738,927

; CURRENT FILING DATE: 2003-12-16

; PRIOR APPLICATION NUMBER: JP 2002-365675

; PRIOR FILING DATE: 2002-12-17

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 777

; TYPE: DNA

; ORGANISM: Manihot esculenta Crantz

US-10-738-927-1

Query Match 22.5%; Score 205.4; DB 21; Length 777;

Best Local Similarity 55.3%; Pred. No. 7.2e-54;

Matches 420; Conservative 0; Mismatches 336; Indels 3; Gaps 1;

Qy 77 CACTTCGTGTAGTTACAAACGCTTATCATGAGCGTGGATCTGTGTACAAAGCTCAAGCC 136

Db 13 CATTTTGTTCTGATTCATACCAATTTGCCAATGGTGTGATGGATTTGGCATTAAGCTCAACCA 72

Qy 137 CTCCTTGAATCAGCGCGCCACCGGTTACTGCTGTGAACTCGCCGCTTCGGGATCGAC 196

Db 73 GCCCTTGAGAGAGCTGGCCACAAAGTCACTGCACTGGACATGGCAGCGCGCATTTGAC 132

Qy 197 CCAGACCAATCCAGCGCGTGTAAACCGTCGACGATACTCCAAACCGTTGATCGAAACC 256

Db 133 CCAAGGCAAAATTTGAGCAGATTAATTTGATGAATACTCTGAACCCCTTATTGACTTTC 192

Qy 257 CTCAAATCTTCCAGAGAACGAAGAGTAAATCTGTTGGATTTCAGCTTCGAGGCGATC 316

Db 193 TTGGAGAACTCCCTCAAGGGGAAAAGGTCATCTTTGTTGGTGTGAGAGCTGTGACGGGCTG 252

Qy 317 AACATGGCTTCGCCCGCAGACATATTTCCGGCGAAGATTAAGGTTCTTTGTGTTCTCTCAAC 376

Db 253 AATATTGCTATTGCTGCTGATAGATACGTTGACAAAATTTGCAGCTGGTGTGTTTCCACAAT 312

Qy 377 GCCTTCTTCCCGACACACACCCACGTCCTTCTCAGGTTCTGGACAAGTATATGGAGATG 436

Db 313 TCCCTTATTGCCAGACACCGCTTATAGCCCATCTTACACTGTGGAAGAGCTTTTGGAGTCG 372

Qy 437 CTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGG---GACGATG 493

Db 373 TTTCTGACTGGAGAGACACAGATATTTTAGTTTCACTAATATCACTGGAGAGACAAAT 432

Qy 494 AGTTTATTGAAGATGGACCAAAATTCATGAAGGACGTCCTTTTACAAAATTTGCCATA 553

Db 433 ACAACAATGAAGCTGGGCTTCGTACTTCTGAGGAAAAATTTATTTACCAAAATGCACATGAT 492

Qy 554 GAGGATTACAGCTGGCAAAAATTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTA 613

Db 493 GGGGAATATGAACTGGCAAAAATTTGTAATGAGGAAGGGATCACTGTTCAAAATGTTTGG 552

Qy 614 TCAGAGAAAGAAAGTTTAGCGAGGAAGGATATGTTTCGGTGCACAGTTCACGTAATG 673

Db 553 GCTCAGAGACCGNAGTTTACCGNAAAAGGTTACGGATCAATTAAGAAAGTTTATATTGG 612

Qy 674 AGTAGTGAAGACAAGCCATCCCCTGCGATTTCATTGTTGGATGATGATATTTCAAC 733

Db 613 ACCGATCAAGACAATAATATTTTACCAGACTTTTCAAGCGCTGGCAATTTGCAAACTACAAA 672

Qy 734 GTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGTGATGCTCTCCAAACCCCAA 793

Db 673 CCAGACAAGGTTTATCAGGTTTCAAGGTGGAGATCATAGCTCCAGCTTACAAAACTGAG 732

Qy 794 AAACCTCTTTGACTCTCTCTCTGCTATTTGCCACCGATTAT 832

Db 733 GAGGTAGCTCATATTCTCCAAAGAGGTGGCTGATGCATAT 771

Search completed: September 12, 2005, 12:48:37

Job time : 2023.05 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 00:50:21 ; Search time 2907.77 Seconds  
(without alignments)  
11938.557 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagcgtattctcaacaa.....ttttccgggcaacttctc 912

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	827	90.7	865	CNS0926U	Arabidops
2	662	72.6	667	AV822206	Arabidops
3	624	68.4	625	AV785311	Arabidops
4	603.4	66.2	641	CB258477	Arabidops
5	557	61.1	572	AV822640	Arabidops
6	533.8	58.5	537	AV783452	Arabidops
7	434.8	47.7	441	AV824147	Arabidops
8	421	45.2	421	BP613009	Arabidops
9	418.6	45.9	480	AA04853	Arabidops
10	397.4	43.6	425	BP642196	Arabidops
11	391.2	42.9	426	AV790585	Arabidops
12	374.2	41.0	390	AV820476	Arabidops
13	366	40.1	366	AV554161	Arabidops
14	359.8	39.5	405	BP642613	Arabidops
15	357.8	39.2	361	AI933246	Arabidops
16	357.6	39.2	382	BP646852	Arabidops
17	298.4	32.7	335	AV545429	Arabidops
18	282.8	31.0	351	BP604502	Arabidops
19	259.2	28.4	797	CF837566	Arabidops
20	247	27.1	308	AV559909	Arabidops
21	246.8	27.1	797	CB290370	Arabidops
22	246.8	27.1	804	CB293837	Arabidops
23	246.8	27.1	821	CB290604	Arabidops
24	246.8	27.1	865	CN190989	Arabidops

RESULT 1

CNS0926U

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA complete sequence from clone GSLTSL1782C11 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Journal

Comment

Annotation

Unpublished

2 (bases 1 to 865)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life technologies (a division of invitrogen)

full-length libraries construction : Temple G

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1..865

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

FEATURES

source

ALIGNMENTS

CNS0926U 865 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA complete sequence from clone  
GSLTSL1782C11 of Silique of strain col-0 of Arabidopsis thaliana  
(thale cress).

Accession BX833804

Version BX833804.1 GI:42455725

Keywords HTC; GSLT cDNA.

Source Arabidopsis thaliana

Organism Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

SpERMophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 865)

Castelli V., Aury J.M., Jaillon O., Wincker P., Clepet C.,

Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V.,

Temple G., Caboche M., Weissenbach J., and Salanoubat M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 865)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life technologies (a division of invitrogen)

full-length libraries construction : Temple G

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1..865

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

```

/clone="GSLTSL1782C11"
/tissue_type="Silique"
/plasmid="PCMVSPORT_6"
1..865
/gene="At5g10300"

ORIGIN
Query Match          90.7%; Score 827; DB 3; Length 865;
Best Local Similarity 99.9%; Pred. No. 9.5e-242;
Matches 838; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 75 ATCACTTCGTGTAGTTCACAACGCTTATCATGGAGCTTGGATCGGTACAAAGCTCAAGC 134
Db 1 ATCACTTCGTGTAGTTCACAACGCTTATCATGGAGCTTGGATCGGTACAAAGCTCAAGC 60

Qy 135 CCTCTTTGAATCAGCCGCGCACCGCGTTACTGTGTGCGAACTCGCCGCTCCGGGATCG 194
Db 61 CCTCTTTGAATCAGCCGCGCACCGCGTTACTGTGTGCGAACTCGCCGCTCCGGGATCG 120

Qy 195 ACCACGACCAATCCAGGCGGTTGAAACCGTGCAGCAATATCTCMAACCGTTGATCGAAA 254
Db 121 ACCACGACCAATCCAGGCGGTTGAAACCGTGCAGCAATATCTCCAAACCGTTGATCGAAA 180

Qy 255 CCTCAAAATCTCTCCAGAGAACGAGGTAATTTCTGTTGGATTTCAGCTTCGGAGGCA 314
Db 181 CCTCAAAATCTCTCCAGAGAACGAGGTAATTTCTGTTGGATTTCAGCTTCGGAGGCA 240

Qy 315 TCAACATCGCTCTCGCCGCGCACATATTTCCGGCGAAGATTAAAGTTCTTGTGTTCTPCA 374
Db 241 TCAACATCGCTCTCGCCGCGCACATATTTCCGGCGAAGATTAAAGTTCTTGTGTTCTPCA 300

Qy 375 AGCCTTCTTGCCCGACACAA-CCACAGTGCCTTCTCAGTTCTGGACAAAGTATATGAG 433
Db 301 AGCCTTCTTGCCCGACACAA-CCACAGTGCCTTCTCAGTTCTGGACAAAGTATATGAG 360

Qy 434 ATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATG 493
Db 361 ATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATG 420

Qy 494 AGTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATGTGCCATA 553
Db 421 AGTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATGTGCCATA 480

Qy 554 GAGGATTACGAGCTGGCAAAATGTTCATAGGCAAGGTCATTTTTTACAGAGGATCTA 613
Db 481 GAGGATTACGAGCTGGCAAAATGTTCATAGGCAAGGTCATTTTTTACAGAGGATCTA 540

Qy 614 TCAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTGCGTCAACGAGTTTACGTAATG 673
Db 541 TCAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTGCGTCAACGAGTTTACGTAATG 600

Qy 674 AGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTGATAATTTCAAC 733
Db 601 AGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTGATAATTTCAAC 660

Qy 734 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCATATGGTGATGCTCTCCAAACCCCAA 793
Db 661 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCATATGGTGATGCTCTCCAAACCCCAA 720

Qy 794 AAATCTTTTGACTCTCTCTGCTATGCGACCGGATATATGTAATAATCTTAAAGTCGT 853
Db 721 AAATCTTTTGACTCTCTCTGCTATGCGACCGGATATATGTAATAATCTTAAAGTCGT 780

Qy 854 TTTACTTTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTTTCCGGGCAACTTTTCATC 912
Db 781 TTTACTTTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTTTCCGGGCAACTTTTCATC 839

RESULT 2
AV822206 667 bp mRNA linear EST 01-APR-2002
LOCUS AV822206 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-02-P20 5',
DEFINITION mRNA sequence.

ACCESSION AV822206 GI:19864240
VERSION AV822206.1
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 667)
AUTHORS Seki,M., Sarusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Nakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.
FEATURES
Location/Qualifiers
source
1..667
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-02-P20"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"

ORIGIN
Query Match 72.6%; Score 662; DB 1; Length 667;
Best Local Similarity 99.7%; Pred. No. 3.5e-191;
Matches 662; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATCAAGCTTATCTCAACACGAAAGACCTTACCAAAAGAAAGAAAGCTTACCGAGACA 60
Db 2 ATCAAGCTTATCTCAACACGAAAGACCTTACCAAAAGAAAGAAAGCTTACCGAGACA 61

Qy 61 TATGAGAGGAAACATCATTCTGTTAGTTTCAACAGCTTATCATGAGGCTCGATCTG 120
Db 62 TATGAGAGGAAACATCATTCTGTTAGTTTCAACAGCTTATCATGAGGCTCGATCTG 121

Qy 121 GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGCAACTCGC 180
Db 122 GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGCAACTCGC 181

Qy 181 CGCTTCGGGATCGACCCACGACCAATCCAGCCGTTGAACCCGTCGACGAATACTCCAA 240
Db 182 CGCTTCGGGATCGACCCACGACCAATCCAGCCGTTGAACCCGTCGACGAATACTCCAA 241

Qy 241 ACCGTTGATCAAAACCCCTCAATCTCTTCCAGAGAACGAGAGGTAATCTTGGTTGGATT 300
Db 242 ACCGTTGATCAAAACCCCTCAATCTCTTCCAGAGAACGAGAGGTAATCTTGGTTGGATT 301

Qy 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGGAAAGATTAAGGT 360
Db 302 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGGAAAGATTAAGGT 361

Qy 361 TCTTGTGTTCTCAACGCTTCTTGGCCGACACAAACCCAGCTGCTTCTCAGTTCTCGA 420
Db 362 TCTTGTGTTCTCAACGCTTCTTGGCCGACACAAACCCAGCTGCTTCTCAGTTCTCGA 421

Qy 421 CAAGTATATGAGATGCTCGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACAAAG 480

```



```

|||||
422 CAAGTATATCGAGATCCCTGGAGGTTTGGAGATTCTGAGTTTTCATCTCATGAAACAAG 481
QY
481 AAATGGGAGCATGAGTTTATGAAGATGGGACCAAAATTCATGAGGCGACGTTCTTACCA 540
Db
482 AAATGGGAGCATGAGTTTATGAAGATGGGACCAAAATTCATGAGGCGACGTTCTTACCA 541
QY
541 AAATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTTTT 600
Db
542 AAATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTTTT 601
QY
601 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAAAGATATGTTCCGTCGCAACG 660
Db
602 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAAAGATATGTTTCNGTGCACG 661
QY
661 AGTT 664
Db
662 AGTT 665

```

```

RESULT 3
AV785311/c
LOCUS
DEFINITION
AV785311 RAFL6 Arabidopsis thaliana cDNA clone RAFL06-14-F12 3',
mRNA sequence.

```

```

ACCESSION
AV785311.1 GI:19804101
VERSION
EST
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

```

```

REFERENCE
AUTHORS
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.

```

```

TITLE
Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL
Unpublished (2002)
COMMENT
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

```

```

An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for
further details.

```

```

FEATURES
source
1..625
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL06-14-F12"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL6"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

```

```

ORIGIN
Query Match 68.4%; Score 624; DB 1; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.5e-179;
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 AAGAGTATTCGTTGGATTTCAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACA 338
|||||

```

```

Db
625 AAGAGTATTCGTTGGATTTCAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACA 566
QY
339 TATTTCCGCGCAAGATTAAAGTTCTTGTGTTCTCAACGCTTCTTGTCCCGCACAAACC 398
Db
565 TATTTCCGCGCAAGATTAAAGTTCTTGTGTTCTCAACGCTTCTTGTCCCGCACAAACC 506
QY
399 ACGTGCTTCTCACGTTCTGGACAAGTATATGGAGATCGCTGGAGGTTTGGAGATTGTG 458
Db
505 ACGTGCTTCTCACGTTCTGGACAAGTATATGGAGATCGCTGGAGGTTTGGAGATTGTG 446
QY
459 AGTTTTCATCATGCAAAAACAAGAAATGGGACGATGAGTTTATGAAGATGGGACCAAAAT 518
Db
445 AGTTTTCATCTCATGAAAACAAGAAATGGGACGATGAGTTTATGAAGATGGGACCAAAAT 386
QY
519 TCATGAAGGCGACGCTCTTTACCAAAATCTCCCATAGAGGATTTACGAGCTGGCAAAATGCT 578
Db
385 TCATGAAGGCGACGCTCTTTACCAAAATGTTCCCATAGAGGATTTACGAGCTGGCAAAATGCT 326
QY
579 TGCATAGGCAAGGGTCATTTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTAGCGAGG 638
Db
325 TGCATAGGCAAGGGTCATTTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTAGCGAGG 266
QY
639 AAGGATATGTTTCGTTGCAACGAGTTTACGTAATGAGTAGTAGTGAAGACAAAGCCATCCCCT 698
Db
265 AAGGATATGTTTCGTTGCAACGAGTTTACGTAATGAGTAGTAGTGAAGACAAAGCCATCCCCT 206
QY
699 GCGATTTTCATTCGTTGGAGTATGATTAATTTCAAGCTCTCGAAAGCTTACGAGATCGATG 758
Db
205 GCGATTTTCATTCGTTGGAGTATGATTAATTTCAAGCTCTCGAAAGCTTACGAGATCGATG 146
QY
759 GCGAGATCATCATGTTGATGCTCTCCAAACCCCAAAAACCTCTTTGACTCTCTCTCTGCTA 818
Db
145 GCGAGATCATCATGTTGATGCTCTCCAAACCCCAAAAACCTCTTTGACTCTCTCTCTGCTA 86
QY
819 TTGCCACCGATTATGTAATTAATCTTAAGTCCGTTTACTTTTTTCTCATCGTTACTTAA 878
Db
85 TTGCCACCGATTATGTAATTAATCTTAAGTCCGTTTACTTTTTTCTCATCGTTACTTAA 26
QY
879 TAAACAACAAACCCCTTTTTCGCGGC 902
Db
25 TAAACAACAAACCCCTTTTTCGCGGC 2

```

```

RESULT 4
CB258477
LOCUS
DEFINITION
CB258477 641 bp mRNA linear EST 06-NOV-2003
05-E012734-014-002-J01-T7R MP12-ADIS-014 Arabidopsis thaliana cDNA
clone MP12p771J012Q 5-PRIME, mRNA sequence.

```

```

ACCESSION
CB258477
VERSION
CB258477.1 GI:32883250
KEYWORDS
EST
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

```

```

REFERENCE
AUTHORS
Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weishaar,B.
TITLE
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL
Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE
22683290
PUBMED
12799357
COMMENT
Contact: Weishaar B
ADIS DNA core facility at MP12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishna@mpiz-koeln.mpg.de
Insert Length: 641 Std Error: 0.00
Plate: 2 row: J column: 01
Seq primer: T7R; CTAATACGACTCTACTATAGGA.

```

```
FEATURES
  source
    Location/Qualifiers
      1..641
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /culivar="Cape Verde Island (Cvi)"
        /db_xref="GABI:590641"
        /db_xref="taxon:3702"
        /clone="MP12p771J012Q"
        /tissue_type="whole plant"
        /dev_stage="adult plant, mixed stresses"
        /lab_host="E. coli TOP10"
        /clone_lib="MP12-ADIS-014"
        /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Cvi; six weeks old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses, (1) at 4 Grad C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a foreceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 Grad C in the light/JUV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation:
T7-Sali-CCACGGCTCCG-Sprime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."
      66.2%; Score 603.4; DB 6; Length 641;
      Best Local Similarity 99.8%; Pred. No. 3.2e-173;
      Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 AGAAGCTTAAACCAAGAAACGCTACGGAGACATATGGAGGAAACATCACTTCGT 84
DB 37 AGAAGCTTAAACCAAGAAACGCTACGGAGACATATGGAGGAAACATCACTTCGT 96

QY 85 GTTAGTTACAAACGCTTATCATGAGCCCTGGATCTGTGACAGCTCAAGCCCTCTTGA 144
DB 97 GTTAGTTACAAACGCTTATCATGAGCCCTGGATCTGTGACAGCTCAAGCCCTCTTGA 156

QY 145 ATCAGCGGCCACCGCGTTACTGTGTGAACTCGCGCGCTCCGGGATCGACCCAGACC 204
DB 157 ATCAGCGGCCACCGCGTTACTGTGTGAACTCGCGCGCTCCGGGATCGACCCAGACC 216

QY 205 AATCCAGCGCGTTGAAACCGTCGACGATCTCCAAACCGTTGTATCGAAACCTCAAAATC 264
DB 217 AATCCAGCGCGTTGAAACCGTCGACGATCTCCAAACCGTTGTATCGAAACCTCAAAATC 276

QY 265 TCTTCCAGAACGAAAGGTAATTCGTGTTGGATTTCAGCTTCGGAGGCATCAACATCGC 324
DB 277 TCTTCCAGAACGAAAGGTAATTCGTGTTGGATTTCAGCTTCGGAGGCATCAACATCGC 336

QY 325 TCTCGCGCCGACATATTTCCGGCGAAGATTAAAGTTCTTGTGTTCTTCAAGCCCTTCTT 384
DB 337 TCTCGCGCCGACATATTTCCGGCGAAGATTAAAGTTCTTGTGTTCTTCAAGCCCTTCTT 396

QY 385 GCCCGACACACCCAGTCTTCTCAGCTTCGACCAAGTATATGGAGATGTCCTGGAGG 444
DB 397 GCCCGACACACCCAGTCTTCTCAGCTTCGACCAAGTATATGGAGATGTCCTGGAGG 456

QY 445 TTTGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATGAA 504
DB 457 TTTGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATGAA 516

QY 505 GATGGGACCAAAATTTCATGAAGCGAGGTCGTTTACCAAAATTTGTCCTCATAGGATTACGA 564

Db 517 GATGGGACCAAAATTTCATGAAGCGAGGTCGTTTACCAAAATTTGTCCTCATAGGATTACGA 576
QY 565 GCTGCGCAAAAATGTTGCATAGGCAAGGTCATTTTTTTCACAGAGGATCTATCAAAAGAAGA 624
DB 577 GCTGCGCAAAAATGTTGCATAGGCAAGGTCATTTTTTTCACAGAGGATCTATCAAAAGAAGA 636
QY 625 AAGT 629
DB 637 AAGT 641

RESULT 5
AV822640 572 bp mRNA linear EST 01-APR-2002
LOCUS AV822640 RAF15 Arabidopsis thaliana cDNA clone RAFL05-09-N03 5', mRNA sequence.
ACCESSION AV822640.1 GI:19864692
VERSION AV822640
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 572)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.
FEATURES
  source
    Location/Qualifiers
      1..572
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /clone="RAFL05-09-N03"
        /dev_stage="rosette plants"
        /lab_host="SOLR"
        /clone_lib="RAFL5"
        /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated (1,2,5,10,24 hr)"

ORIGIN
  Query Match 61.1%; Score 557; DB 1; Length 572;
  Best Local Similarity 99.5%; Pred. No. 5.2e-159;
  Matches 568; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCAGCTATTCTCAACAGCAAGAACCTTACCAAAAGAAACGCTCTACGGAGACA 60
DB 2 ATCAGCTATTCTCAACAGCAAGAACCTTACCAAAAGAAACGCTCTACGGAGACA 61

QY 61 TATGAGAGGAAACATCACTTCGTGTTAGTTTCAACAGCTTATCATGAGGCTCGATCTG 120
DB 62 TATGAGAGGAAACATCACTTCGTGTTAGTTTCAACAGCTTATCATGAGGCTCGATCTG 121

QY 121 GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGTCGAACCTGC 180
DB 122 GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGTCGAACCTGC 181
```



```

further details.
FEATURES             Location/Qualifiers
  source             1..441
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /db_xref="taxon:3702"
                    /clone="RAFL06-14-F12"
                    /dev_stage="plants at various developmental stages from
                    germination to mature seeds"
                    /lab_host="DH10B"
                    /clone_lib="RAFL6"
                    /note="Site 1: SstI; Site 2: XhoI; subjected to
                    dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                    hr) treatments"

ORIGIN
Query Match       47.7%; Score 434.8; DB 1; Length 441;
Best Local Similarity 99.1%; Pred. No. 1.4e-121;
Matches 436; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATFCAGCTATTCTCAACAGCAAGAAACCTTTAACCAAAAGAAACCTCTACGGAGACA 60
DB 2 ATFCAGCTATTCTCAACAGCAAGAAACCTTTAACCAAAAGAAACCTCTACGGAGACA 61
QY 61 TATGAGAGGAACATCATCTTGTGTAGTTTCACAAACGCTTATCATGGAGCCTGGATCTG 120
DB 62 TATGAGAGGAACATCATCTTGTGTAGTTTCACAAACGCTTATCATGGAGCCTGGATCTG 121
QY 121 GTACAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGTGCGAACTCGC 180
DB 122 GTACAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGTGCGAACTCGC 181
QY 181 CGCCTCCGGGATCGACCCACGACCAATCCAGCCGCTTGAAACCGTCGACGAATCTCCA 240
DB 182 CGCCTCCGGGATCGACCCACGACCAATCCAGCCGCTTGAAACCGTCGACGAATCTCCA 241
QY 241 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGAACGAGGTAATCTGGTTGGATT 300
DB 242 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGAACGAGGTAATCTGGTTGGATT 301
QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCAGCATATTTCCGGGGAAGATTAAAGT 360
DB 302 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCAGCATATTTCCGGGGAAGATTAAAGT 361
QY 361 TCTTGTTGTTCTCAACGCTTCTTGCCGACACAAACCCACGCGCTTCTCAGCTTCGGA 420
DB 362 TCTTGTTGTTCTCAACGCTTCTTGCCGACACAAACCCACGCGCTTCTCAGCTTCGGA 421
QY 421 CAAAGTATATGAGATGCGCTG 440
DB 422 CAAAGTTCGGANATGCGCTG 441

RESULT 8
BP613009/c         421 bp  mRNA  linear  EST 26-JUN-2004
LOCUS             BP613009 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-91-K04 3',
DEFINITION        mRNA sequence.
ACCESSION         BP613009.1 GI:49264191
VERSION           BP613009
KEYWORDS          EST.
SOURCE            Arabidopsis thaliana (thale cress)
ORGANISM          Arabidopsis thaliana
REFERENCE         1 (bases 1 to 421)
AUTHORS           Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
                   Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
                   Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
                   Functional annotation of a full-length Arabidopsis cDNA collection
                   Science 296 (5565), 141-145 (2002)
TITLE             Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL           Science 296 (5565), 141-145 (2002)

```

```

MEDLINE            21932900
PUBMED            11910074
COMMENT           Contact: Motoaki Seki
                   Plant Functional Genomics Research Group
                   RIKEN Genomic Sciences Center
                   3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
                   Tel: 81-298-36-4359
                   Fax: 81-298-36-9060
                   Email: msek@rtc.riken.go.jp
                   reversed clone; Please visit our web site
                   (http://pfigweb.gsc.riken.go.jp/) for further details.
FEATURES           Location/Qualifiers
  source           1..421
                  /organism="Arabidopsis thaliana"
                  /mol_type="mRNA"
                  /db_xref="taxon:3702"
                  /clone="RAFL16-91-K04"
                  /lab_host="DH10B"
                  /clone_lib="RAFL16"
                  /note="Site_1: BamHI; Site_2: SalI; dark-grown"

ORIGIN
Query Match       46.2%; Score 421; DB 5; Length 421;
Best Local Similarity 100.0%; Pred. No. 2.4e-117;
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 492 TGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGCCCA 551
DB 421 TGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGCCCA 362
QY 552 TAGAGGATTACGAGCTGCAAAAATGTTGCATAGCAAGGTCATTTTTCACAGAGATC 611
DB 361 TAGAGGATTACGAGCTGCAAAAATGTTGCATAGCAAGGTCATTTTTCACAGAGATC 302
QY 612 TATCAAAAGAAAGAAAGTTTAGCGAGGAAGATATGTTTCGGTCAACAGAGTTTACGTA 671
DB 301 TATCAAAAGAAAGAAAGTTTAGCGAGGAAGATATGTTTCGGTCAACAGAGTTTACGTA 242
QY 672 TGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTCA 731
DB 241 TGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTCA 182
QY 732 ACCTCTCGAAGTCTACGAGATCGATGCGGAGATCACATGTCATGCTCTCCAAACCCC 791
DB 181 ACCTCTCGAAGTCTACGAGATCGATGCGGAGATCACATGTCATGCTCTCCAAACCCC 122
QY 792 AAAAATCTTTTGACTCTCTCTCTCTATTCGCCAGGATTAATATGTAATAATCTTAAAGTCC 851
DB 121 AAAAATCTTTTGACTCTCTCTCTCTATTCGCCAGGATTAATATGTAATAATCTTAAAGTCC 62
QY 852 GTTTTACTTTTCTCATCGTTACTTAATAAACAACCCCTTTTTCGGGGCAACTTCAT 911
DB 61 GTTTTACTTTTCTCATCGTTACTTAATAAACAACCCCTTTTTCGGGGCAACTTCAT 2

QY 912 C 912
DB 1 C 1

RESULT 9
AA404853          28251 Lambda-PRL2 Arabidopsis thaliana cDNA clone 143M11XP 3', mRNA
LOCUS             AA404853
DEFINITION        sequence.
ACCESSION         AA404853
VERSION           AA404853.1 GI:2062871
KEYWORDS          EST.
SOURCE            Arabidopsis thaliana (thale cress)
ORGANISM          Arabidopsis thaliana
REFERENCE         1 (bases 1 to 480)
AUTHORS           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
                   1 (bases 1 to 480)

```

**AUTHORS** Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H.,  
 McIntosh, L., Ohlroge, J., Rakheh, N., Somerville, S., Thomas, M.,  
 Rezel, E. and Somerville, C.  
**TITLE** Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
**JOURNAL** Plant Physiol. 106, 1241-1255 (1994)  
**PUBMED** 95148729  
**COMMENT** 7846151  
 Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, Mi  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tcn@bm.cl.msu.edu  
 Seq primer: M13 -21 dye primer.  
 Location/Qualifiers  
 1. 480  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="143M11Xp"  
 /clone\_lib="Lambda-PRL2"  
 /note="vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

```

RESULT 10
BP642196/c 425 bp mRNA linear EST 27-JUN-2004
LOCUS BP642196 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-57-N06 3',
DEFINITION mRNA sequence.
ACCESSION BP642196.1 GI:49293666
VERSION BP642196
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 425)
Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakajima,M., Enju,A., Akiyama,K., Ono,Y., Muramatsu,M.,
Hayaishizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shingawa,A. and Shinozaki,K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
JOURNAL Science 296
MEDLINE 21932900
PUBMED 11910074
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@riken.go.jp
reversed clone; Please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.
FEATURES
source Location/Qualifiers
1..425
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL19-57-N06"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL19"
/note="Site_1: BamHI; Site_2: SalI; Subtraction Library"
ORIGIN
Query Match 43.6%; Score 397.4; DB 5; Length 425;
Best Local Similarity 98.5%; Pred. No. 4.le-110; Indels 0; Gaps 0;
Matches 401; Conservative 0; Mismatches 6;

Qy 506 ATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGCCATAGAGGATTACGAG 565
Db 425 ATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTTCCATAGAGGATTACGAG 366
Qy 566 CTGGCAAAATTTGCATAGCGAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAA 625
Db 365 CTGGCAAAATTTTGCATAGCGAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAA 306
Qy 626 AAGTTTAGCAGGAAGGATATGTTTCGGTCGACAGGTTTACGTAATGAGTAGTGAACAC 685
Db 305 AAGTTTAGCAGGAAGGATATGTTTCGGTCGACAGGTTTACCTAATGAGTAGTGAACAC 246
Qy 686 AAAGCCATCCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCAACGTCGAAAGTC 745
Db 245 AAAGCCATCCCCTGCGATTTTCATTCGTTGGATGATTGATAATTTCAACGTCGAAAGTC 186
Qy 746 TAGAGATCGATGCGGAGATCATCGTGATGCTCTCCAAACCCCAAAACTCTTTGAC 805
Db 185 TAGAGATCGATGCGGAGATCATCGTGATGCTCTCCAAACCCCAAAACTCTTTGAC 126
Qy 806 TCTCTCTCTGCTATTGCGACCGATTATGTAATTAATCTTTAAGTCCGTTTTTACTTTTTTC 865
Db 125 TCTCTCTCTGCTATTGCGACCGATTATGTAATTAATCTTTAAGTCCGTTTTTACTTTTTTC 66
Qy 866 TCATCGTTACTAATAAACAACCCCTTTTTTTCGGGGCAACTTTTCATC 912

```

```

Db
|||||
65 TCATCGTTACTAATAAAACAACCCCTTTTCCGGCCAACTTTCATC 19

RESULT 11
AV790585/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 426)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for
further details.
FEATURES
source
1..426
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF106-89-G15"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL6"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
ORIGIN
Query Match 42.9%; Score 391.2; DB 1; Length 426;
Best Local Similarity 99.0%; Pred. No. 3.3e-108;
Matches 404; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 506 ATGGGACCAAAATTCATGAAGCAGTCCTTTACCAAAATTTGCCATAGAGGATTACGAG 565
Db 426 ATGGGACCAAAATTCATGAAGCAGTCCTTTACCAAAATTTGCCATAGAGGATTACGAG 367
QY 566 CTGGCAAAAATCTTGCAATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAGAA 625
Db 366 CTGGCAAAAATCTTGCAATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAGAA 307
QY 626 AAGTTTAGCGAGGAAGGATATCGTTTCGGTGCAACGAGTTTACGTAATGAGTAGTGAAGAC 685
Db 306 AAGTTTAGCGAGGAAGGATATCGTTTCGGTGCAACGAGTTTACATATGAGTAGTGAAGAC 247
QY 686 AAGCCATCCCTGCGATTTCATTGTTGGATGA-TTGNATATTTCAACGTTCTCGAAGT 744
Db 246 AAGCCATCCCTGCGATTTCATTGTTGGATGAGTTGATATTTCAACGTTCTCGAAGT 187
QY 745 CTCAGAGATCGATGGCGGAGATCAATGGTGTATGCTCTCCAAACCCCAAAACTCTTTGA 804
|||||

Db
186 GTACGAGATCGATGGCGGAGATCACATGGTGTGCTCTCCAAACCCCAAAACTCTTTGA 127
QY 805 CTCTCTCTCTGCTATTGCGCACCGGATTATATGTAATAATCTTAAGTCGGTTTACTTTTT 864
|||||
126 CTCTCTCTCTGCTATTGCGCACCGGATTATATGTAATAATCTTAAGTCGGTTTACTTTTT 67
QY 865 CTATCGTTTACTAATAAAACAACCCCTTTTTCGGGCAACTTTTCATC 912
|||||
66 CTATCGTTTACTAATAAAACAACCCCTTTTTCGGGCAACTTTTCATC 19

RESULT 12
AV820476/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 390)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
FEATURES
source
1..390
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL11-10-111"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL11"
/note="Site 1: BamHI; Site 2: SalI; subjected to various
treatments (dehydration, cold, high salt, ABA, heat and
UV). Dark-grown plants"
ORIGIN
Query Match 41.0%; Score 374.2; DB 1; Length 390;
Best Local Similarity 99.0%; Pred. No. 5.2e-103;
Matches 387; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 522 TGAAGGCACGCTTTTACCAAAATTTGCCATAGAGGATTACGAGTCGCAAAAATGTTGC 581
Db 390 TGAAGGCACGCTTTTACC-AAATGGTCCCATAGAGGATTACGATCTGCGCAAAATGTTGC 332
QY 582 ATAGCGAAGGTCATTTTTCACAGAGGATCTCTCAAGAGAAAGAACTTTAGCGGGAAG 641
|||||
331 ATAGCGAAGGTCATTTTTCACAGAGGATCTCTCAAGAGAAAGAACTTTAGCGGGAAG 272
QY 642 GATATGGTTCGGTGCACAGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGC 701

```

```

Db      271  GATATGTTGGTGCACACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGGG 212
Qy      702  ATTTTCATTGGTTCGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGG 761
Db      211  ATTTTCATTGGTTCGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGG 152
Qy      762  GAGATCACATGGTGTGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTG 821
Db      151  GAGATCACATGGTGTGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTG 92
Qy      822  CGACCGATTATGTTAATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAA 881
Db      91   CCACCGATTATGTTAATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAA 32
Qy      882  AACAAACCCCTTTTCCGGGCAACTTTTCATC 912
Db      31   AACAAACCCCTTTTCCGGGCAACTTTTCATC 1

```

```

RESULT 13
AV554161
LOCUS   AV554161 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION
cDNA clone RZ82h03R 5', mRNA sequence.
ACCESSION
AV554161
VERSION
AV554161.1 GI:8725574
KEYWORDS
EST
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 366)
AUTHORS
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL
DNA Res. 7 (3), 175-180 (2000)
MEDLINE
20363093
PUBMED
10907847
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..366
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/scf_type="Columbia"
/db_xref="taxon:3702"
/clone="RZ82h03R"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

```

```

ORIGIN
Query Match 40.1%; Score 366; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.7e-100;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 72 AACATCACTTCGTTAGTTAGTTCACAGCTTATCATGGAGCTTGGATCTGTGTAAGCTCA 131
Db 1 AACATCACTTCGTTAGTTAGTTCACAGCTTATCATGGAGCTTGGATCTGTGTAAGCTCA 60
Qy 132 AGCCCTCTCTTGAATCAGCGGCGCCAGCGGTACTCTGTGAACTCGCGCGCTCCGGGA 191
Db 61 AGCCCTCTCTTGAATCAGCGGCGCCAGCGGTACTCTGTGAACTCGCGCGCTCCGGGA 120
Qy 192 TCGACCCACGACCAATCCAGGCGGTTGAACCGTCGACGAATACTCCAAACCGTTGATCG 251

```

```

Db      121  TCGACCCACGACCAATCCAGGCCGTTGAACCGTCGACGAATATCTCCAAACCGTTGATCG 180
Qy      252  AAACCCCTCAAAATCTCTTCCAGAGAACGAAGAGTAATTTCTGTTGGATTTCAGCTTCGGAG 311
Db      181  AAACCCCTCAAAATCTCTTCCAGAGAACGAAGAGTAATTTCTGTTGGATTTCAGCTTCGGAG 240
Qy      312  GCATCAACATCGCTCTCGCCGCGACATATTTCCGGCCGAAGATTAAAGTTTCTTGTTCC 371
Db      241  GCATCAACATCGCTCTCGCCGCGACATATTTCCGGCCGAAGATTAAAGTTTCTTGTTCC 300
Qy      372  TCAACGCTTCTTTCGCCGACACCAACCCAGCGCTTCTCACGTTCTGCAAGATATATGG 431
Db      301  TCAACGCTTCTTTCGCCGACACCAACCCAGCGCTTCTCACGTTCTGCAAGATATATGG 360
Qy      432  AGATGC 437
Db      361  AGATGC 366

RESULT 14
BP642613/c
LOCUS   BP642613 RAPL19 Arabidopsis thaliana cDNA clone RAFL19-59-D05 3',
DEFINITION
mRNA sequence.
ACCESSION
BP642613
VERSION
BP642613.1 GI:49294083
KEYWORDS
EST
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 405)
AUTHORS
Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
TITLE
Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL
Science 296 (5565), 141-145 (2002)
MEDLINE
21932900
PUBMED
11910074
COMMENT
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
reversed clone; Please visit our web site
(http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES
source
1..405
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL19-59-D05"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL19"
/note="Site_1: BamHI; Site_2: SalI; Subtraction Library"

```

```

ORIGIN
Query Match 39.5%; Score 359.8; DB 5; Length 405;
Best Local Similarity 99.0%; Pred. No. 1.4e-98;
Matches 383; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
Qy 528 CACGCTTTTACAAAATTGTCCCATAGA-GGATTACGAGCTGGCAAAATGTTGCATAGG 586
Db 405 CACGCTTTTACAAAATTGTCCCATAGAGGATTACGAGCTGCCAAAATGTTGCATAGC 346
Qy 587 CAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTTAGCGAAGAGGATAT 646
Db 345 CAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTTAGCGAAGAGGATAT 286

```

QY	647	GGTTCGGTGCAACGAGTTTACGTAATAGTAGTGAAGAACAAGCCATCCCCTGCGATTTC	706
Db	285	GGTTCGGTGCAACGAGTTTACGTAATAGTAGTGAAGAACAAGCCATCCCCTGCGATTTC	226
QY	707	ATTTCGTTGGATGATGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGGCGGAGAT	766
Db	225	ATTTCGTTGGATGATGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGGCGGAGAT	166
QY	767	CACATGGTGATGCTCTCCAAACCCCAAAACCTCTTGTGACTCTCTCTGCTA-TTGCCAC	825
Db	165	CACATGGTGATGCTCTCCAAACCCCAAAACCTCTTGTGACTCTCTCTGCTAGTTGCCAC	106
QY	826	CGATTATATGTAATAATCTTAAGTCGTTTACTTTTTCATCGTTTACTTAATAAACA	885
Db	105	CGATTATATGTAATAATCTTAAGTCGTTTACTTTTTCATCGTTTACTTAATAAACA	46
QY	886	RAACCCCTTTTCCGGGCACTTTCATC	912
Db	45	AAACCCCTTTTCCGGGCACTTTCATC	19
RESULT 15			
AI993246			
LOCUS			
DEFINITION			
AI993246			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY			
Db			
QY			

Db	61	CATCACTTCGTGTAGTTTACAAAGCTTATCATCGAGCCTGGATCTGGTACAAAGCTCAAG	120
QY	134	CCCCTCCTTGAATCAGCCGCGCCACCGGTTACTGCTGTGAACTCGCGCGCTCCGGGATC	193
Db	121	CCCCTCCTTGAATCAGCCGCGCCACCGGTTACTGCTGTGAACTCGAGCGCTCCGGGATC	180
QY	194	GACCCAGCAATCCAGGCGGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAA	253
Db	181	GACCCAGCAATCCAGGCGGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAA	240
QY	254	ACCTTCAAAATCTCTTCAGAGAACGAAAGAGTAACTCTGGTTGGATTTCAGCTTCGGAGGC	313
Db	241	ACCTTCAAAATCTCTTCAGAGAACGAAAGAGTAACTCTGGTTGGATTTCAGCTTCGGAGGC	300
QY	314	ATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAAGATTAAAGGTTCTTTGTTCCTC	373
Db	301	ATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAAGATTAAAGGTTCTTTGTTCCTC	360
QY	374	A	374
Db	361	A	361

Search completed: September 12, 2005, 06:31:55  
Job time : 2912.77 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 22:50:50 ; Search time 499.745 Seconds  
(without alignments)  
10803.121 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagctattctcaacaa.....ttttcgggcaactttcatc 912

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	912	100.0	984	8	ADA26534	Ada26534 Arabidops
2	910.4	99.8	1056	3	AAC47895	Aac47895 Arabidops
3	887.8	97.3	909	3	AAC47800	Aac47800 Arabidops
4	481.4	52.8	483	10	ADE82083	Ades82083 Arabidops
5	243	26.6	918	8	ADA26530	Ada26530 Arabidops
6	241.6	26.5	792	6	ABZ13515	Abz13515 Arabidops
7	241.6	26.5	816	3	AAC42344	Aac42344 Arabidops
8	232.6	25.5	1079	8	ADA26498	Ada26498 Tobacco s
9	222.6	24.4	1152	8	ADA26532	Ada26532 Arabidops
10	221.6	24.3	1091	2	AAT36351	Aat36351 Hevea bra
11	219.2	24.0	817	2	AAT86824	Aat86824 Hevea bra
12	218.8	24.0	792	3	AAC43133	Aac43133 Arabidops
13	218.8	24.0	792	6	ABZ12412	Abz12412 Arabidops
14	218.8	24.0	792	12	ADN74620	Adn74620 Thale cre
15	214.2	23.5	792	6	ABZ13746	Abz13746 Arabidops
16	214.2	23.5	864	3	AAC47519	Aac47519 Arabidops
17	214.2	23.5	870	3	AAC34411	Aac34411 Arabidops
18	214.2	23.5	950	8	ADA26531	Ada26531 Arabidops
19	205.4	22.5	777	12	ADQ16305	Adq16305 Nucleotid
20	201.8	22.1	823	3	AAC42529	Aac42529 Arabidops

21	201.8	22.1	928	8	ADA26538	Ada26538 Arabidops
22	199.6	21.9	1109	8	ADA26533	Ada26533 Arabidops
23	198.6	21.8	777	12	ADQ16306	Adq16306 Nucleotid
24	194.4	21.3	771	6	ABZ13473	Abz13473 Arabidops
25	193.4	21.2	770	8	ADA67827	Ada67827 Arabidops
26	192.2	21.1	834	8	ADA26537	Ada26537 Arabidops
27	185.2	20.3	783	8	ADA26536	Ada26536 Arabidops
28	178.6	19.6	795	8	ADA70991	Ada70991 Rice gene
29	176.8	19.4	1046	8	ADA26535	Ada26535 Arabidops
30	172.6	18.9	828	3	AAC41733	Aac41733 Arabidops
31	172.6	18.9	903	3	AAC34141	Aac34141 Arabidops
32	172.6	18.9	908	8	ADA26539	Ada26539 Arabidops
33	141.2	15.5	422	3	AAC43356	Aac43356 Arabidops
34	136.2	14.9	984	3	AAC49239	Aac49239 Arabidops
35	122.6	13.4	1325	3	AAC47916	Aac47916 Arabidops
36	115.6	12.7	783	8	ADA70766	Ada70766 Rice gene
37	96.6	10.6	978	8	ADA69955	Ada69955 Rice gene
38	91.4	10.0	860	8	ADA71305	Ada71305 Rice gene
39	91.2	10.0	1041	8	ADA26546	Ada26546 Arabidops
C	91	10.0	541	10	ABX57306	Abx57306 Arabidops
	72.4	7.9	910	3	AAC35440	Aac35440 Arabidops
42	72.4	7.9	1516	8	ADA26547	Ada26547 Arabidops
43	71.4	7.8	1815	3	AAC44497	Aac44497 Zea mays
44	71.2	7.8	720	8	ADA70729	Ada70729 Rice gene
45	68.4	7.5	789	6	ABZ12505	Abz12505 Arabidops

## ALIGNMENTS

RESULT 1  
ADA26534  
ID ADA26534 standard; DNA; 984 BP.

AC ADA26534;

DT 20-NOV-2003 (first entry)

DE Arabidopsis thaliana AtSB2L5 gene homologous to tobacco SABP2 gene.

XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

PN WO2003016551-A2.

PD 27-FEB-2003.

PF 16-AUG-2002; 2002WO-US026312.

PR 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Klessig DF, Kumar D;

DR WPI; 2003-278578/27.

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.

XX Claim 65; Page; 98pp; English.

PS The invention relates to a novel isolated 1079 bp salicylic acid-binding  
protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
protein, sequences that specifically hybridize with the nucleic acid, the  
complement of the nucleic acid or a natural allelic variant of the  
sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
conferring increased resistance to disease in plants, or for screening  
modulators of SABP2, which confer increased or enhanced resistance of

plants to disease. The SAMP2 or SAMP2 homologue nucleic acid is also useful for producing transgenic plants with increased resistance to disease. The SAMP2 nucleic acid, SAMP2, or its homologue, is also useful for studying the molecular mechanisms responsible for the modulation of salicylic acid-mediated disease resistance in plants. This sequence represents the gene encoding the Arabidopsis thaliana SAMP2 ALSB2L5 which has homology to the gene (ADA26498) encoding the tobacco SAMP2 protein. (Note: this sequence is not given in the specification but is derived from the Genbank accession number given by the inventors).

Sequence 984 BP; 291 A; 230 C; 200 G; 263 T; 0 U; 0 Other; XX

Query Match	100.0%;	Score 912;	DB 8;	Length 984;
Best Local Similarity	100.0%;	Pred. No. 2e-286;		
Matches 912.	Concentrations	0.	Matches 0.	Indels

Qy	1	ATCACGCTATTCTCAAAACGAAAGAACTTTAACCAAAAGAAAGAAAGCGTCTACGGAGACA	60
Db	1	ATCACGCTATTCTCAAAACGAAAGAAAGCGTTTAAACCAAAAGAAAGAAAGCGTCTACGGAGACA	60
Qy	61	TATGGAGAGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGGAGCCTGGATCTG	120
Db	61	TATGGAGAGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGGAGCCTGGATCTG	120
Qy	121	GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGGCGACCGCGTTACTGCTGTGAACTCGC	180
Db	121	GTACAAAGCTCAAGCCCTCTTGAATCAGCGCGGCGACCGCGTTACTGCTGTGAACTCGC	180
Qy	181	CGCCTCCGGGATCGACCCAGCACCACCAATCCAGCGCGTTGAACCGTCCAGCAATACTCCAA	240
Db	181	CGCCTCCGGGATCGACCCAGCACCACCAATCCAGCGCGTTGAACCGTCCAGCAATACTCCAA	240
Qy	241	ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAAAGAAAGGTAATTCCTGGTTGGATT	300
Db	241	ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAAAGAAAGGTAATTCCTGGTTGGATT	300
Qy	301	CAGCTTCGGAGGATCGAATCGCTCTCGCGCGCGACATATTTCCGGCGGAAGATTAAAGT	360
Db	301	CAGCTTCGGAGGATCGAATCGCTCTCGCGCGCGACATATTTCCGGCGGAAGATTAAAGT	360
Qy	361	TCCTGTGTCTCTCAACGGCTCTTTCGCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA	420
Db	361	TCCTGTGTCTCTCAACGGCTCTTTCGCCGACACAAACCCACGTCGCTTCTCACGTTCTGGA	420
Qy	421	CAAGTATATGGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACAAG	480
Db	421	CAAGTATATGGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACAAG	480
Qy	481	AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGCAGCGCTTTTACCA	540
Db	481	AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGCAGCGCTTTTACCA	540
Qy	541	AAATTGTCCATAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT	600
Db	541	AAATTGTCCATAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT	600
Qy	601	CACAGAGGATCTATCAAGAAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTCCGTTGCAACG	660
Db	601	CACAGAGGATCTATCAAGAAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTCCGTTGCAACG	660
Qy	661	AGTTTACGTAACGATAGTAGTAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGAT	720
Db	661	AGTTTACGTAACGATAGTAGTAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGAT	720
Qy	721	TGATAATTTCAACGCTCTCGAAAGTCTTACGAGATCGATGCGGAGATCACATGGTGATGCT	780
Db	721	TGATAATTTCAACGCTCTCGAAAGTCTTACGAGATCGATGCGGAGATCACATGGTGATGCT	780
Qy	781	CTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCGATTATATGTAATA	840
Db	781	CTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCGATTATATGTAATA	840
Qy	841	ATCTTAAGTCCGTTTACTTTTTTCTCATCGTTTACTATAAAACCAAAACCCCTTTTCCGG	900

PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158023P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159323P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		

Query Match

Best Local Similarity

Matches 911; Conservative

Score 910.4; DB 3;

Length 1056;

Pred. No. 6.8e-286;

Mismatches 1;

Indels 0;

Gaps 0;

QY

1

ATCAGCTATTCTCAACAAAGAACCTTTAACCAAAAAGAAACGTTTACGGAGACA

60

Db

128

ATCAGCTATTCTCAACAAAGAACCTTTAACCAAAAAGAAACGTTTACGGAGACA

187

QY 61 TATGAGAGGAACATCACTTCGTAGTTAGTTCACAAACGCTTATCATGAGCGCTGGATCTG 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 188 TATGAGAGGAACATCACTTCGTAGTTAGTTCACAAACGCTTATCATGAGCGCTGGATCTG 247  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 GTACAAGCTCAAGCCCTCCCTTGAATCAGCGCGCCACCGCTTACTGCTGCAACTCGC 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 248 GTACAAGCTCAAGCCCTCCCTTGAATCAGCGCGCCACCGCTTACTGCTGCAACTCGC 307  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 CGCTCCCGGATCGACCCACGACCAATCAAGCGCTTGAACCGCTGACGCAATATCCAA 240  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 308 CGCTCCCGGATCGACCCACGACCAATCAAGCGCTTGAACCGCTGACGCAATATCCAA 367  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 ACCGTTGATCGAAACCCCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 300  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 368 ACCGTTGATCGAAACCCCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 427  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGACATATTTCCGGCGAAGATTAAGGT 360  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 428 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGACATATTTCCGGCGAAGATTAAGGT 487  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 TCTTGTTCTCAACGCTTCTTCCCGACACAACCCACGTCCTTCTCAGTTCTGGA 420  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 488 TCTTGTTCTCAACGCTTCTTCCCGACACAACCCACGTCCTTCTCAGTTCTGGA 547  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 CAACTATATGAGATCGCTCGGAGGTTTGGAGATTTGGAGATTTTCATCTCATGAACAAG 480  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 548 CAACTATATGAGATCGCTCGGAGGTTTGGAGATTTGGAGATTTTCATCTCATGAACAAG 607  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 AAATGGGACGATGAGTTTATCAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 540  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 608 AAATGGGACGATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 667  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 AAATGTCCATPAGAGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 600  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 668 AAATGTCCATPAGAGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 727  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGTTCCGGTCAACG 660  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 728 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGTTCCGGTCAACG 787  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGATGAT 720  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 788 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGATGAT 847  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 721 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCAATGCTGATGCT 780  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 848 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCAATGCTGATGCT 907  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 781 CTCCAAACCCGAAACCTTTTGACTCTCTCTGCTATTCACCGGATATATGTAATA 840  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 908 CTCCAAACCCGAAACCTTTTGACTCTCTCTGCTATTCACCGGATATATGTAATA 967  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 841 ATCTTAAGTCGGTTTACTTTTCTCATCGTTTACTAATAAAACAAACCCCTTTTCCGG 900  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 968 ATCTTAAGTCGGTTTACTTTTCTCATCGTTTACTAATAAAACAAACCCCTTTTCCGG 1027  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 901 GCAACTTTCATC 912  
DB |||||||||||  
QY 1028 GCAACTTTCATC 1039  
DB |||||||||||

RESULT 3  
AAC47800  
ID AAC47800 standard; DNA; 909 BP.  
AC AAC47800;  
XX AAC47800;  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55168.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;

```
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-014342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 97.3%; Score 887.8; DB 3; Length 909;
Beat Local Similarity 99.7%; Pred. No. 1.5e-278;
Matches 900; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCAGCTATTCTCAACAAGCAAGAACCTTAAACCAAAAGAAACCTTACCGGAGACA 60
DB 7 ATAAGCTATTCTCAACAAGCAAGAACCTTAAACCAAAAGAAACCTTACCGGAGACA 66

QY 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAACCTTATCATGAGAGCTCGATCTG 120
DB 67 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAACCTTATCATGAGAGCTCGATCTG 126

QY 121 GTACAAGCTCAAGCCCTCTTGTGAATACGCGGCCACCGCGTTACTGTGCGAACTCGC 180
DB 127 GTACAAGCTCAAGCCCTCTTGTGAATACGCGGCCACCGCGTTACTGTGCGAACTCGC 186

QY 181 CGCCTCCGGATCGACCCGACCAATCCAGCCGTTGAACCGTCGACGAACTCTCCAA 240
DB 187 CGCCTCCGGATCGACCCGACCAATCCAGCCGTTGAACCGTCGACGAACTCTCCAA 246

QY 241 ACCGTTGATGAAACCCCTCAAAATCTTCCAGAGAACGAGAGGTAATTTCTGGTTGGATT 300
DB 247 ACCGTTGATGAAACCCCTCAAAATCTTCCAGAGAACGAGAGGTAATTTCTGGTTGGATT 306

QY 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 360
DB 307 CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 366
```

```
QY 361 TCTTGTGTTCTCAACGCTTCTTCCCGCACACACCAACCGTGCCTTCTCAGTTCGGA 420
Db 367 TCTTGTGTTCTCAACGCTTCTTCCCGCACACCAACCGTGCCTTCTCAGTTCGGA 426
QY 421 CAAGTATATGGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG 480
Db 427 CAAGTATATGGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG 486
QY 481 AAATGGGACGATGAGTATTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 540
Db 487 AAATGGGACGATGAGTATTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 546
QY 541 AAATTGTCCCATAGAGGATTACGAGCTGCGCAAAATGTTGCATAGGCAAGGTCATTTTT 600
Db 547 AAATTGTCCCATAGAGGATTACGAGCTGCGCAAAATGTTGCATAGGCAAGGTCATTTTT 606
QY 601 CACAGAGGATCTATCAAGAAAGAAAGTTTATAGCGAGGAAGGATATGTTGGTGCAACG 660
Db 607 CACAGAGGATCTATCAAGAAAGAAAGTTTATAGCGAGGAAGGATATGTTGGTGCAACG 666
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 720
Db 667 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 726
QY 721 TGATTAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCATGCGTGATGCT 780
Db 727 TGATTAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCATGCGTGATGCT 786
QY 781 CTCCTAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATGTCACCGATTATATGTAATA 840
Db 787 CTCCTAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATGTCACCGATTATATGTAATA 846
QY 841 ATCTTT-AAGTCGCGTTTACTTTTTTCTCATCGTTACTAATAAAACCAACCCCTTTTCGG 899
Db 847 ATCTTAAAGTCGTTTACTTTTTTCTCATCGTTACTAATAAAACCAACCCCTTTTCGG 906
QY 900 GGC 902
Db 907 GGC 909

RESULT 4
ADE82083
ID ADE82083 standard; cDNA; 483 BP.
AC ADE82083;
XX
XX
DT 29-JAN-2004 (first entry)
DE
DE Arabidopsis thaliana expressed polynucleotide seq id 854.
XX
KW genetically modified organism; transgenic organism; plant;
KW inhibitor testing; activator testing; modifier testing; fungicide;
KW insecticide; genetic function; genetic regulation; cellular metabolism;
KW gene; ss.
XX
XX Arabidopsis thaliana.
XX
XX US2003115639-A1.
XX
XX 19-JUN-2003.
XX
XX 26-JAN-2001; 2001US-00770961.
XX
XX 27-JAN-2000; 2000US-0178466P.
XX
XX (GORLACH J.
XX (ANYI Y.
XX (HAMILTON C M.
XX (PRICE J L.
XX (RAIN/ RAINES T M.
XX (YU Y.
XX (RAME/ RAMEKA J G.
```

```
PA (PAGE/ PAGE A.
PA (MATH/ MATHW A V.
PA (LEDF/ LEDFORD B L.
PA (WOES/ WOESSNER J P.
PA (HAAS/ HAAS W D.
PA (GARC/ GARCIA C A.
PA (KRIC/ KRICKER M.
PA (SLAT/ SLATER T.
PA (DAVI/ DAVIS K R.
PA (ALLE/ ALLEN K.
PA (HOFF/ HOFFMAN N.
PA (HURB/ HURBAN P.
XX GORLACH J, AN Y, HAMILTON CM, PRICE JL, RAINES TM, YU Y;
PI RAMEKA JG, PAGE A, MATHW AV, LEDFORD BL, WOESSNER JP, HAAS WD;
PI GARCIA CA, KRICKER M, SLATER T, DAVIS KR, ALLEN K, HOFFMAN N;
PI HURBAN P;
XX WPI; 2003-810930/76.
XX
XX Novel Arabidopsis thaliana nucleic acids useful for generating
XX genetically modified transgenic organisms, for screening biologically
XX active agents such as fungicides, insecticides.
XX
XX Claim 1; SEQ ID NO 854; 44pp; English.
XX
XX The invention describes a nucleic acid (I) comprising a sequence capable
XX of hybridising under stringent conditions to any one of 999 fully defined
XX Arabidopsis thaliana sequences (I) as given in specification e.g., 360,
XX 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a
XX hybridisation probe to complementary molecules in a cDNA library. (I) is
XX also useful for generating genetically modified and transgenic organisms,
XX usually plant cells and plants. A protein encoded by (I) is useful in
XX screening assays to determine the effect of candidate inhibitors,
XX activators or modifiers of the gene product. The protein is also useful
XX for screening biologically active agents e.g., fungicides and
XX insecticides. A genetically modified cell, comprising an exogenous
XX nucleic acid, where the nucleic acid comprises transcription regulatory
XX sequences operably linked to a sequence capable of hybridising under
XX stringent conditions to (I) is useful in the study of genetic function
XX and regulation, for alteration of the cellular metabolism and for
XX screening compounds that may affect the biological function of the gene
XX or gene product. This sequence represents an Arabidopsis thaliana
XX polynucleotide of the invention.
XX
XX Sequence 483 BP; 129 A; 127 C; 112 G; 115 T; 0 U; 0 Other;
SQ
Query Match 52.8%; Score 481.4; DB 10; Length 483;
Best Local Similarity 99.8%; Pred. No. 5.3e-146;
Matches 482; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 GAACTTTAAACCAAAAGAAAACGCTACGGGAGACATATGGAGAGAAACATCCTTCGTG 85
Db 1 GAACTTTAAACCAAAAGAAAACGCTACGGGAGACATATGGAGAGAAACATCCTTCGTG 60
QY 86 TTAGTTCAACACGCTTATCATGAGCGCTGGATCTGGTACAAGCTCAAGCCCTCCTTGAA 145
Db 61 TTAGTTCAACACGCTTATCATGAGCGCTGGATCTGGTACAAGCTCAAGCCCTCCTTGAA 120
QY 146 TCAGCGGCCACCGGTTACTGCTGAACTCCGGCTCCGGATCGACCCAGCACCA 205
Db 121 TCAGCGGCCACCGGTTACTGCTGAACTCCGGCTCCGGATCGACCCAGCACCA 180
QY 206 ATCCAGGCGGTTGAAACCGTCGACGATATCTCCAAACCGTTGATCGAAATCT 265
Db 181 ATCCAGGCGGTTGAAACCGTCGACGATATCTCCAAACCGTTGATCGAAATCT 240
QY 266 CTTCCAGAGAACGAGAGGTAATTCGTTGGATTTCAGCTTCGAGGATCAACATCGCT 325
Db 241 CTTTCAGAGAACGAGAGGTAATTCGTTGGATTTCAGCTTCGAGGATCAACATCGCT 300
QY 326 CTGCGCGCGGACATATTTCCGCGCAAGATTAAAGGTTCTTGTTCTCTCAACGCCCTTCTTG 385
```



XX 28-FEB-2002.  
PD  
XX PF  
XX 24-AUG-2001; 2001WO-US026685.  
XX  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Harper JF, Kreps J, Wang X, Zhu T;  
PI WPI; 2002-304127/34.  
XX  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
XX Claim 144; SEQ ID NO 1320; 577pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
XX Sequence 792 BP; 217 A; 167 C; 196 G; 212 T; 0 U; 0 Other;  
SQ  
Query Match 26.5%; Score 241.6; DB 6; Length 792;  
Best Local Similarity 59.6%; Pred. No. 1.5e-67;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCATTGCTGTAGTTTACACACGCTTATCATGAGCGTGGATCTGTGACA 125  
DB 14 AGAGGAAACAAATTTGCTAGTGTACATGCTGTGCGCATGCGGCTGCTGTGACA 73  
QY 126 AGCTCAAGCCCTCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCAACTCGCGGCT 185  
DB 74 AGGTTAAGCGGCTGTAGAGCGGTGGCCACCGGTTACTGCTGTGCAACTGCTGTGCT 133  
QY 186 CCGGGATCGACCCA---CGACCAATCCAGCGGCTTGAACCGGTGACGAACTCTCAAC 242  
DB 134 CCGGAATAGACACACGAGGTGCTGACTGACATCCCATCCGACATCGCAACAATCTCGGAGC 193  
QY 243 CGTTGATCGAACCCTCAATCTCTCCAGAGAACGAGAGCTAATCTGTTGGATTCA 302  
DB 194 CATTGACGAAGTCTCTGACCTCATTCGCAATGATGAAAGGTGTGCTGTTGTCACA 253  
QY 303 GCTTCGGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTTC 362  
DB 254 GCTTTGTTGCTTGAACCTTAGCCATAGCCATGMAAAGTTTCCGMAAATCTGTGCG 313  
QY 363 TTGTGTTCTCAACGGCTTCTTGCCGCGACACACCGCATGCTTCTCAAGTTCTGGACA 422  
DB 314 CTGTATTCTTGTGCTGCTTTCATGCGCGACACCGAACATCTCACCATCTTCTGCTTGGACA 373  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAA 476  
DB 374 AGTTTGGAGACACATGCTCTCAAGAACATGATGGGACCGGAATTCGAACTTATGTTT 433  
QY 477 CAAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGACGCTTTT 536  
DB 434 CAGACAATTCGGACTGAGTATGTT---TTTATAGCCCTGACTTTCATGAAGTTGGGTCTCT 490  
QY 537 ACCAAATTTGTCCCATAGAGGATTACGACTGGCAAAATGTTGATAGGCAAGGTCAT 596

DB 491 ACCAGCTTCTCCAGTTGAGGATCTTGAATGGGATTACTTTTATGAGGCCAGATCGT 550  
QY 597 TTTTTCACAGAGGATCTATCAAAAGAAAGAAAGTTTATAGCGAGGAAAGGATATGTTGCGTGC 656  
DB 551 TATTTATTAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGTCTGTTTC 610  
QY 657 AACGAGTTTACGTAATGAGTAGTAGTGAAGACAAAGCATCCCTCGGATTTTCAATTCGTTGGA 716  
DB 611 CTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATTTCCAGAAGAACGCCAGAGATGGA 670  
QY 717 TGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGA 776  
DB 671 TGATTGATAATTTTCCCGTGAATTTAGTGAATGGAGATGGAGAGACAGATCATATGCCAA 730  
QY 777 TGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTTATTCGCCACCGATTATATGT 836  
DB 731 TGTCTGCAAGCCTCAGCACTCAGTGATTAATCTTCTGAAATTCGGGACAAATTCGTTT 790  
QY 837 AA 838  
DB 791 AA 792  
RESULT 7  
AAC42344  
ID AAC42344 standard; DNA; 816 BP.  
XX AAC42344;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.



PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139917P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

```
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      26.5%; Score 241.6; DB 3; Length 816;
Best Local Similarity 59.6%; Pred. No. 1.5e-67;
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;

QY 66 AGAGGAACATCAGTCGCTGGTTAGTTTCAACGCTTATCATGAGCGCTGGATCTGGTACA 125
DB 38 AGAGGAACAACATTTGTTACTAGTACATGCTGTCGCATGGCGTGGTCTGGTACA 97

QY 126 AGCTCAAGCCCTCTCTTGAATCAGCGCGCCACCGGTTACTCTCTGCGAACTCGCGGCT 185
DB 98 AGGTTAAGCGCTGCTAGAGGCGGTGGGCCACCGGTAACTGCTGTGACTTAGCTGCT 157

QY 186 CCGGATCGAACCCA---CGACCAATCCAGGCGCTTGAACCGTTCGACGAATCTCCAAC 242
DB 158 CCGGAATAGACACAACGAGGTGATCACTGACATCCCAATGCCACATCGCAACAATACTCGGAGC 217

QY 243 CGTTGATCGAAACCTCAATCTCTTCCAGAGAACGAAGAGTAATCTGTGTTGATTC 302
DB 218 CATTGACGAAGCTCTGACCTCATTTGCCAATGATGAAGAAGTTGTGCTGTTGTCACA 277

QY 303 GCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTTC 362
DB 278 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAGAAGTTTCCGGAANAATCTGTGCG 337

QY 363 TTGTGTTCTCAACGCTTCTTTCGGCAGACAAACCGACGTCCTTCTCAAGTTCTGGACA 422
DB 338 CTGTATTTCTGACTGCTTTCATGCGGACACCGCACTCAACATCTTCGTTCTGGACA 397

QY 423 AGTAT-----ATGGAGATGCTGGAGTTTGGGAGATGTGAGTTTTCATCTCATGAA 476
DB 398 AGTTTGGGAAGCAACATGCTCAAGAAGCATGTGATGGGCACCGAAATCGAACCTTATGTT 457

QY 477 CAAGAAATGGGACGATGAGTTTATTAAGATGGGACCAAAATTCATGAAGGCACGCTTTT 536
DB 458 CAGACAAATCCGAGCTGAGTATGTT---TTTAGCCCTGACTTCATGAAGTTGGTCTCT 514

QY 537 ACCAAAATGTCCCATAGAGGATACAGCTCGGAAAAAATGTTGCATAGGCAAGGTCAT 596
DB 515 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGCCAGGATCGT 574

QY 597 TTTTCACAGAGGATCTATCAAGAGAAAGATTTTAGCGAGGAAGATATGTTTCGGTGC 656
DB 575 TATTTATTAACGATTTATCGAAGATGAANAATCTTCGGATGAAGATATGGTCTGTTTC 634

QY 657 AACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCGATTTCAATTCGTTGA 716
DB 635 CTCGAGTTTTCATAGTGTGTNAGGACAAAGCAATTCAGAGAACCGCCAGAGATGA 694

QY 717 TGATTGATAATTTCAACGCTCGAAGTCTACAGATCGATGCGGAGATCATGTTGA 776
DB 695 TGATTGATAATTTTCGGTGAATTTAGTGTAGGATGAGATGAGGAGACAGATCATATGCCAA 754

QY 777 TGCTCTCCAAACCCCAAACTTTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 836
DB 755 TGTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAATTCGGGAATTCGGGAATTCGTTT 814

QY 837 AA 838
DB 815 AA 816
```

RESULT 8  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX

```
AC ADA26498;
XX
DT 20-NOV-2003 (first entry)
XX
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.
XX
KW salicylic acid-binding protein; SABP2; disease resistance; plant;
modulator; tobacco; ds; gene.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
CDS 42..824
FT /*tag= a
FT /product= "salicylic acid binding protein SABP2"
XX
PN WO2003016551-A2.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US026312.
XX
PR 16-AUG-2001; 2001US-0312863P.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Klessig DF, Kumar D;
XX
DR WPI; 2003-278578/27.
XX
P-PSDB; ADA26499.
XX
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for
producing transgenic plants with increased resistance to disease, or for
screening SABP2 modulators that confer enhanced resistance of plants to
disease.
XX
PS Claim 1; Fig 4; 98pp; English.
XX
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding
protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid
protein, sequences that specifically hybridize with the nucleic acid, the
complement of the nucleic acid or a natural allelic variant of the
sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for
confering increased resistance to disease in plants, or for screening
modulators of SABP2, which confer increased or enhanced resistance of
plants to disease. The SABP2 or SABP2 homologue nucleic acid is also
useful for producing transgenic plants with increased resistance to
disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful
for studying the molecular mechanisms responsible for the modulation of
salicylic acid-mediated disease resistance in plants. This sequence
CC represents the gene encoding the SABP2 protein from tobacco plants.
XX
SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Query Match      25.5%; Score 232.6; DB 8; Length 1079;
Best Local Similarity 56.3%; Pred. No. 1.5e-64;
Matches 458; Conservative 0; Mismatches 349; Indels 6; Gaps 1;

QY 33 AACCAAAAGAAAACGCTCTACGGAGACATATGGAGAGGAAACATCATTCTGTGTAGTTC 92
DB 13 AAAAGAACTAACAGGCATAAAATTCAAATGAAGGAAGGAAACACATTTGTTTAGTAC 72

QY 93 ACAAGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGCCCTCTCTTGAATCAGCG 152
DB 73 ATGTTGCATGCCATGGAGGTTGGAGTTGGTACAGCTAAAGCCACTGCTAGAGCTGAG 132

QY 153 GCCACCGCTTACTGTCTCGAACTCGCGGCTCCGGGATCGACCCAGCAATCCAGG 212
DB 133 GCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGAGAAATAGAGG 192

QY 213 CCGTTGAAACCGTTCGACGAATCTCCAAACCGTTGATCGAAACCCCTCAATCTCTTCAG 272
DB 193 AGCTTCGCACACTTTATGATTATATCTTTGCCAATTGATGAGTGTGATGAATCTCTTTTCAG 252
```

Qy	273	AGAACGAAGAGGTAAATCTGGTTGGATTACGCTTCGGAGGCATCAACATCGCTCTGCGC	332
Db	253	CAGATGAGAAAGGTATATTAGTGGGCA TAGTCTTGGTGTATGAATTTGGGACTTGCTA	312
Qy	333	CCGACATATTTCCGGCGGAAGATTAAAGTTCTTGTGTTCTCAACGGCTTCTTGC	392
Db	313	TGGAAGATATCCACAAGAGATCTATGCTGCTGTTTCTTGTGCTGTTTTCATG	372
Qy	393	CAACCCAGCTGCTTCTCAGGTTCTGGACAAGTATATGGAGATGCTGGAG-----GTT	446
Db	373	CTGTTCAACATCTCCTTTGTTTGGAAACAGTAAATGAGCGGACGCCAGCGGAGAA	432
Qy	447	TGGGAGATTTGAGTGTTCATCTCATGAAAACAAGAAATGGGACGATGAGTTATTGAAGA	506
Db	433	GTTGGATACTCAGTTTTTACATATGTTCCCTCGAGAGGCCATGACATCCATGTTTT	492
Qy	507	TGGGACCAAAATTCATAGGAGCGCTCTTTACCAAAATTTGCCATAGAGGATTCAGAGC	566
Db	493	TTGGCCAAAGTTCTTGGCTACAAGCTCTACCAGCTATGCTCTCTGAGGATCTTGAT	552
Qy	567	TGGCAAAATGTTGCATAGGCAAGGTCATTTTTACAGAGGATCTATCAAAAGAAGAA	626
Db	553	TAGCATCATATGTTGGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCA	612
Qy	627	AGTTTAGCGGAGGATATGTTGCTCGGTGCAACGAGTTTACGTAATGAGTACTGAAGCA	686
Db	613	ATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGTCATGAGGATA	672
Qy	687	AAGCCATCCCCTGCCATTTTCAATTGGAATGATTGAATAATTTCAACGTCCTCGAAGTCT	746
Db	673	AAGGATACGAGAGAAATTCAGCGATGGCAATTTGAACAAATGGTGTCTAGAGCA	732
Qy	747	ACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAACCTCTTTGAT	806
Db	733	TAGAGATTAAGGTGCTGATCACATGGCAATGCTATGCGAGCCCCCAAAACCTTTGGCCT	792
Qy	807	CTCTCTCTGCTATTGCCACCGAATTATGTAAT	839
Db	793	CTCTCTTGGAAATTTGCCCATAAATACAACTGAT	825

screening SAMP2 modulators that confer enhanced resistance of plants to disease.

Claim 65; Page; 98pp; English.

The invention relates to a novel isolated 1079 bp salicylic acid-binding protein 2 (SAMP2) nucleic acid molecule encoding a 261 amino acid protein, sequences that specifically hybridize with the nucleic acid, the complement of the nucleic acid or a natural allelic variant of the sequence. The nucleic acid, SAMP2 or a SAMP2 homologue is useful for conferring increased resistance to disease in plants, or for screening modulators of SAMP2, which confer increased or enhanced resistance of plants to disease. The SAMP2 or SAMP2 homologue nucleic acid is also useful for producing transgenic plants with increased resistance to disease. The SAMP2 nucleic acid, SAMP2, or its homologue, is also useful for studying the molecular mechanisms responsible for the modulation of salicylic acid-mediated disease resistance in plants. This sequence represents the gene encoding the Arabidopsis thaliana SAMP2 AtS92L3 which has homology to the gene (ADA26498) encoding the tobacco SAMP2 protein. (Note: this sequence is not given in the specification but is derived from the Genbank accession number given by the inventors).

Sequence 1152 BP; 372 A; 214 C; 253 G; 313 T; 0 U; 0 Other;

Query Match 24.4%; Score 222.6; DB 8; Length 1152;  
Best Local Similarity 56.0%; Pred. No. 3e-61;  
Matches 464: Conservative 0; Mismatches 359; Indels 6;

Qy	16	AACNAGCAAGAACCTTTAA	CCAAAAGAAAACGCTTACGGAGACATAT	TGGAGAGGAAACA	75
Db	93	AAGAAGTCGAGAGACA	AAAGTTTAAAGTAGAGAAGAAAATGAGT	TGAGGAGGAGAGAGAAACA	152
Qy	76	TCACCTTCGTGTAGTTTCA	CAACGCTTATCATGGAGCCTGGATCTCGT	TACAACGCTCAAGCC	135
Db	153	ACACGTCGTTCTAGTACAT	TGGTGTCTGGCCATGGCCCTGGTGTCTG	TACAGAGTTTAAGCC	212
Qy	136	CCTCCTTTGAATACGCGCG	CACCGCTTACTGTGTGTGAACTCGCGCGCT	CTCGGATCGA	195
Db	213	GCAGCTCGAGGCTTCTG	GGCACCGCGTAACCGCGGTAGATCTAG	CTGTGCTCTCGGTATAGA	272
Qy	196	C---CCACGACCAATCAT	CAGGCGGTGAAACCGTCGACGAATACTCT	CAACCGTTCGATCGA	252
Db	273	CATGACACAGTCAATCA	CAGATATATCCACATGCGAACATACTC	CAGAGCCATTCATGCA	332
Qy	253	AACCCCTCAAAATCTCT	CCAGAGAACGAGGTAATTTCTGGTTGSA	TTTTCAGCTTCGGAGG	312
Db	333	GCTAAATGACCTCACTAC	CAAGATGATGAGAAGTTGTCTGTCTGT	CTCATAGCTTAGGAGG	392
Qy	313	CATCAACATCGCTCTCG	CGCGCGACATATTTCCGGCGNAGATTAA	GTTTCTTGTTGTTCTCT	372
Db	393	TTTGAGTTTAGCTATGG	CCCATGGAATGTTTTCCGACCAAAATCT	CTGTCTCTGTCTTTGT	452
Qy	373	CAACGCGCTTCTTGCC	CGACACCAACCCACGTGCTTCTCACCGT	CTTGGACAAGTATATGCA	432
Db	453	GACTGCTATGATGCC	AGACACCAACACTCACCATCCTTCTGT	TGGGATAGCTTAAGAA	512
Qy	433	GATGCCCTGGAGGTTT	GGGAGATTGTGA---GTTTTCATCTCAT	TGAAACAAGAAATGGGAC	489
Db	513	AGAAATCTTACGAGAG	GAATGGTTAGACACCGTGTTTACGAGCG	AGAAACCTGATTTTCC	572
Qy	490	GATGAGTTTATCGAAG	TGGACCAAAATTCATCAAGGCACGTCT	TTTACCAAAATTTGCC	549
Db	573	TACCGAGTTTGGATTT	TTTGACACGAATTCATGGCCAGAAC	CTTTGATTCAGTTGTTCTCC	632
Qy	550	CATAGAGGATTACGAG	CTGGCAAAAATGTTGTCATAGGCAAGG	GTCAATTTTTTTCACAGAGA	609
Db	633	AGTCCAAGATCTTTG	AAATTTGCGCAAAATGTTGGT	TGAGGCAAAACCATTTGATTAAGAAAGA	692
Qy	610	TCTATCAAGAAAGAA	AGTTTACGAGGAGAGGATATGGT	TCGGTGCACACGAGTTTACGT	669
Db	693	TATGGCAGAGAGAA	AGCTTTCAGTGTAGGAAGGATACGG	ATCCCGTTTACACGCTATATTTAT	752
Qy	670	AATGAGTAGTGAAGA	CAAAAGCCATCCCGCTCGAATTTCA	TTCTGTTGGATGATTGATAATTT	729

Db 753 TGTATGCGGAAGGATCTTGTGTACCCGAGATTACAGCGATCGATGATCAGCACTT 812  
QY 730 CAACGCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACC 789  
Db 813 TCCCCCAAAAGAAGTAATGGAGATCAAGAGCGCAGATCATATGCCAATGTTCTCCAAAGCC 872  
QY 790 CCAAAAACCTTTTGACTCTCTCTGCTATTCGCCACCGATTATATGTAA 838  
Db 873 TCAACAACCTATGTGCTCTTCTTGGAGATTGCAAAATAAATATGCTTAA 921

## RESULT 10

AAT36351

ID AAT36351 standard; cDNA; 1091 BP.

XX

AC AAT36351;

XX

DT 29-OCT-1997 (first entry)

XX

DE Hevea brasiliensis S-hydroxynitrilase cDNA.

XX

KW S-hydroxynitrilase; chiral synthesis; aliphatic; aromatic; S-cyanohydrin;  
KW recombinant; ss.

XX

OS Hevea brasiliensis.

XX

FH Key Location/Qualifiers

FT CDS 57..830

FT /\*tag= a

FT /product= "S-hydroxynitrilase"

XX

PN W09703204-A2.

XX

XX 30-JAN-1997.

XX

XX 10-JUL-1996; 96WO-EP003010.

XX

XX 12-JUL-1995; 95AT-00001182.

XX

(STAM ) DSM CHEMIE LINZ GMBH.

XX

PI Hasslacher M, Schall M, Schwab H, Hayn EM, Kohlwein S, Griengl H;

XX

DR WPI; 1997-119058/11.

DR P-PSDB; AAW10593.

XX

PT Hevea brasiliensis S-hydroxynitrilase - and recombinant equivalent of

XX

XX high specific activity, for chiral synthesis of cyano:hydrin(s).

XX

PS Claim 2; Page 2-3; 14pp; German.

XX

CC The present sequence encodes the Hevea brasiliensis S-hydroxynitrilase (S  
CC -HN), which can be used for the chiral synthesis of aliphatic or aromatic  
CC S-cyanohydrins from HCN, or its donor, and ketone. Recombinant S-HN has  
CC higher specific activity than the native enzyme. H. brasiliensis leaves  
CC were homogenised in cold pH 6.5 K phosphate buffer, and the extract  
CC subjected to sequential chromatography on QAE-Sepharose, Phenyl-Sepharose  
CC and BioGel 150 to recover a protein with a specific activity of 19 IU/mg.  
CC A cDNA library was prepared from young H. brasiliensis leaves in Zap  
CC phage, and screened with polyclonal rabbit antiserum raised against the  
CC purified enzyme. The insert in one positive clone was isolated, and  
CC ligated into pBluescript to form pHNL-100. This plasmid, in E. coli SOUR,  
CC expressed a fusion protein with lacZ that was immunoreactive. Sequencing  
CC showed that the insert in this plasmid was not complete, and the missing  
CC 5'-end generated by 2 stage PCR, resulting in plasmid pHNL-101,  
CC containing the entire coding sequence

XX

SQ Sequence 1091 BP; 326 A; 196 C; 239 G; 330 T; 0 U; 0 Other;

Query Match

Best Local Similarity 24.3%; Score 221.6; DB 2; Length 1091;

Matches 431; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

QY 76 TCATCTCGTGTAGTTACAAAGCCTTATATGAGAGCCTTGGATCTGGTACAGCTCAAGCC 135  
Db 68 TCATTTTGTCTTATTCATACCATATGCGCGGTGCATGGATTTGGCAAGCTCAAAACC 127  
QY 136 CCTCTTGAATCAGCCGCGCCACCGCGTTACTGCTGTGGAACCTCGCCGCTCCGGGATCGA 195  
Db 128 CCTCTTGGGCACTTGGCCACAAGGTTACTGCACTGGACCTTGCAAGCAAGCGGGTTGA 187  
QY 196 CCCAGCACAATCCAGCCGTTGAAACCGTCGACGAATATCCAAACCGTTGATCGAAAC 255  
Db 188 CCCAAGGCAAAATTGAGGAGATTGGCTCATTTGATGAGTATTTCTGAACCTTGTGTGACGTT 247  
QY 256 CCTCAATCTCTCCAGAGAACGAGAGGTAATTTCTGTTGCGATTTCAGCTTCGAGGCGAT 315  
Db 248 CTTGGAGGCACTCCCTCCAGGGGAAAGGTAATTTCTGTTGGCGAGAGCTGTGGAGGACT 307  
QY 316 CAACATCGCTCTCGCCGCGGACATATTTCCGCGCGAAGATTAAAGTTCTTGTGTCTCTCAA 375  
Db 308 CAATATAGCAATTGCTGCTGATTAATACTGTGAAAGATTTCAGCTGCTGTTTCCACAA 367  
QY 376 CGCTCTTCTGGCCGACACCAACCGTCGCTTCTCAGCTTCTGGACAAGTATATGGAGAT 435  
Db 368 TTCAGTATTGCCAGACACCGAGCACTGCCATCTTACGTCGTGGATAAGCTCATGGAGGT 427  
QY 436 GCCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGTAATGGACGATGAG 495  
Db 428 GTTTCGCCGACTGGAAAGACACACGTAATTTTACGTACACTAAAGATGCCAAGGAGATAAC 487  
QY 496 TTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTCTTTTACCAAAATTTGTCCCATAGA 555  
Db 488 TGGATTGAACTGGGCTTCACGCTTCTGAGGGAATTTATATACCTTTTCGGTCTCTGA 547  
QY 556 GGATTACGAGCTGGCAAAATTTGTCATAGCAAGGAGGTCATTTTTCACAGAGGATCTATC 615  
Db 548 GGAATATGAACCTGGCGAAGATGTTGACAAAGAGGAGATCATTTTCAAAATATTTTAGC 607  
QY 616 AAAGAAAGAAAGTTTAGCGAGGAGGATATGTTTCGTTGCGTGCACAGTTCAGCTAAATGAG 675  
Db 608 TAAGCGACCATTCCTCACTAAGGAGGTTACGGATCGAATTAAGAAAATTTATGTGTGAC 667  
QY 676 TAGTGAAGACAAAGCCATCCCTCGGATTTTCATTTCTGTTGGATGATTGATAATTTCAACGT 735  
Db 668 CGACCAAGCAAAATATTTTACCTGAATTTTCAACTCTGGCAAAATAGAAAATCTATAAAC 727  
QY 736 CTCGAAAGCTTACGAGATCGATGGCGGAGATCAGATGATGCTCTCTCCAAACCCCAAA 795  
Db 728 AGACAAGGTTTATAAGGTCGAAGGTGGAGATCATAAATTTGCAGCTTACAAAGACTAAGGA 787  
QY 796 ACTCTTTGACTCTCTCTGCTATTTGCCACCGATTATATGTAATAATCTTAAAGTCCGTTT 855  
Db 788 GATCGCTGAAATTTCTCAAGAGGTGGCTGATACCTATTAATTGACTTCTTTGAGGCTTTT 847

## RESULT 11

AAT86824

ID AAT86824 standard; cDNA; 817 BP.

XX

AC AAT86824;

XX

DT 09-JAN-1998 (first entry)

XX

DE Hevea brasiliensis (S)-hydroxynitrilase cDNA.

XX

KW (S)-hydroxynitrilase; production; (S)-cyanohydrin; aldehyde;  
KW hydrogen cyanide; recombinant; ss.

XX

OS Hevea brasiliensis.

XX

FH Key Location/Qualifiers

FT CDS 44..817

FT /\*tag= a

FT /product= "(S)-hydroxynitrilase"